10-1-2012

Discrete arginine topologies guide escape of miniature proteins from early endosomes to the cytoplasm

Jacob S. Appelbaum
Yale University

Discrete arginine topologies guide escape of miniature proteins from early endosomes to the cytoplasm.

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Jacob S. Appelbaum

Dissertation Director: Alanna Schepartz

1 October 2012

Abstract

Discrete arginine topologies guide escape of miniature proteins from early endosomes to the cytoplasm.

Jacob S. Appelbaum

2012

Polypeptides and peptide mimetics sample a wide chemical space with broad potential to modulate cellular function, but their application to cytoplasmic targets is limited because when added to cells their cytosolic concentration remains low. This limitation is due to a diffusion barrier (the plasma membrane) and absence of dedicated import machinery. Highly cationic peptides and proteins sometimes gain cytosolic access, but how they do so is not well understood. Using a small library of cationic miniature proteins, I probe the influence of positive charge number and orientation on the ability the miniature protein to access the cytoplasm. Using a novel assay, I identify a cationic miniature protein, which we called 5.3, that carries a discrete arginine motif and efficiently reaches the cytoplasm. Database searches find that the precise motif identified (an arginine present in positions i, $i+4$, $i+7$, $i+10$, and $i+11$ of an α-helix) is not present in nature, but that similar motifs are present in natural proteins that interact with cellular membranes. Finally, I examine the cellular pathway by which 5.3 reaches the cytoplasm. I find that this miniature protein enters the cell via a dynamin and cholesterol dependent endocytic mechanism and is delivered to Rab5+ early endosomes. In contrast to the shiga-like toxins, and many non-enveloped viruses (which escape to the cytoplasm from the endoplasmic reticulum) as well as other peptides previously identified as 'cell penetrating', only 5.3 escapes from early endosomes. These findings should enable the future dissection of the precise molecular events underlying cytoplasmic access of peptides and proteins, and may illuminate principles for the engineering of peptides and peptidomimetics that access cytoplasmic targets.

Contents

List of Figures

Acknowledgements

Much of the work described in chapters 3 and 4 was recently submitted as a manuscript for publication. My coauthors have provided extensive assistance in the collection of data and in the interpretation of results; for that I am in their debt. This project began with Betsy Smith, a former graduate student, and Douglas Daniels, a former postdoctoral associate in the Schepartz lab. Their research and creativity laid the ground work for these experiments. In addition, the assistance of Justin Holub and Jonathan LaRochelle has been very valuable. Justin and Betsy in particular have contributed results to the figures presented in chapter 3, and I have acknowledged their assistance specifically where their data is presented.

The conception of the project for detecting peptide-MHC complexes came after extensive discussions with Lelia Delamarre and Ona Bloom. This project would not have been possible without their expertise, their patience with me, and their contribution of reagents. I am also indebted to the members of the Cresswell lab, in particular David Stepensky and Ralf Leonhardt for their critical analysis, problem solving skills, and reagent contributions.

The contributions of friends and coworkers has been the foundation of both my success in and enjoyment of science. My classmate Daniel Balkin has contributed significantly in both these regards. He has been the critical ear for preliminary results, a guru with the microscope, a source encouragement, and of many, many plasmids.

I am forever grateful for those who taught me to love science, and I have been lucky to have a large number of teachers in this regard. Begining with my parents, and continuing with my (very special) first grade teacher Aki Kurose, laboratory mentors Leland Hartwell, Linda Buck, Al Singer, Dominic Poccia, and former professors Caroline Goutte, and Pat Williamson, all of whom demonstrated the rewards and fun that are part of the scientific process. As an MD/PhD student I am also indebted to Al Singer for instilling in me the importance of scientific honesty and integrity.

I am very thankful for the hard work and intellectual contributions of my thesis committee, and my advisor. Alanna Schepartz has put together a laboratory that has been remained on the cutting edge for her entire independent career. The ability to consistently inspire her students and garner their success, combined with a track record of creativity is what drew me to her lab. I still find myself excited about future projects I could do, were I only to have more time!

Finally, I must thank my family, and especially my wife, for patience, encouragement, and unconditional love.

Jacob Appelbaum
New Haven, CT
Dec, 2011

Chapter 1

Introduction

Permeability barriers presented by lipid bilayers

Without the ability to separate their contents from the outside world, cells could not be considered 'alive'. The cellular limiting membrane, a phospholipid bilayer, is critical to the concentration of metabolites, generation and use of energy, segregation of genetic material, and sequestration of harmful digestive enzymes or poisons. At the same time, lipid bilayers represent frustrating barriers to the physician or biologist wishing to perturb the inner workings of a cell for the purposes of scientific experiment or therapeutic gain.

The high resistance presented by lipid bilayers is well known in the field of neurobiology, where the increase in conductance due to the opening of single ion channels can be recorded[1]. This is only possible because the membrane (in this case, along with its seal to the recording pipette) prevents the crossing of small ions, resulting in a resistance of $10^9 - 10^{11}\,\Omega$, which, given a thickness of ~5 nM makes the lipid membrane one of the most resistive materials known($\rho \approx 10^{19}\Omega \cdot m$, teflon is ~$10^{22}\Omega \cdot m$, air is ~$10^{16}\Omega \cdot m$).

The efforts of physicians and chemists to identify useful drugs has also been a *de facto* investigation of the permeability barrier presented by cell membranes. While many drugs act on protein receptors on the cell surface, for drugs to act on intracellular targets,

they must reach the cytoplasm. The many examples of bioactive materials, along with the negative examples provided by their inert counterparts, has led to a set of guidelines for the design of small molecules that are readily absorbed and reach the cytoplasm[2]. These guidelines include a limit on molecular size ($\leq$500 Da), polarity (logP > 5), and hydrogen bonding potential ($\leq$ 10 H-bond acceptors, $\leq$5 H bond donors). Thus the limitations of membrane permeability have designated a large fraction of potential chemical compounds as unlikely to be useful.

This limitation has particularly frustrated efforts to manipulate protein protein interactions within cells. These interactions take place over flat surfaces, generally $\geq$600Å^2, often being between 1000-2500Å^2, and thus difficult to target with small molecules. In addition, prediction of key residues within the protein-protein interface is a challenge[3]. While the characteristic physicochemical properties of protein interfaces vary, many contain key polar residues[4]. In fact, within the many residues that form a protein-protein interface, stabilizing binding energy often results from only several residues (termed 'hot spots' [5, 6]) chosen from within a subset of amino acids: phenylalanine, tyrosine, tryptophan, arginine, and isoleucine are over-represented[6]. Small molecules are not well suited to achieve specific binding to flat and sometimes polar surfaces.

Peptides and proteins on the other hand are very well suited to the modulation of protein-protein interactions[7]. By using a peptide that exhibits a well-defined structure in solution, the entropic cost associated with a concerted binding/folding event can be paid 'up front'[8, 9, 10], increasing the magnitude of the free energy of binding. While important advances in the design of small molecule inhibitors based on the crystal structures are being made[11, 12], accurate results are often only achievable as comparisons between structurally similar small molecules[13], and *de novo* identification of high affinity protein ligands has been difficult[13, 12]. The situation for peptides is very different. Protein epitope grafting[14, 15, 16], wherein the interacting side chains of a protein or peptide ligand are grafted onto a well folded α-peptide domain[17, 15, 18, 19, 20, 21, 22] or β-peptide helix[23, 24, 25, 26, 27, 28] enables the facile creation of high affinity ligands.

Peptides with low micromolar affinities can be isolated from small, designed libraries, often containing only a dozen or so members[15]. In the case of naturally encoded α-peptide miniature proteins, ligand affinity can be improved via phage display[29], in the case of non-natural β-peptides, selection can be accomplished from bead-based libraries[30]. If the molecules thus selected could be engineered to reach their cytoplasmic targets, a large selection of rapidly designable active chemical entities would likely become available, enabling scientists and physicians to modulate cytoplasmic protein-protein interactions for the purposes of interrogating function or ameliorating disease.

Those seeking to engineer cytoplasmic access into peptides and proteins may take notice of the fact that pathogens have envolved to achieve the same result for different ends: the debilitation and exploitation of a suitable host. The viral life cycle requires cytoplasmic machinery for the replication of the viral genome[31]. Anthrax and tetanus toxins cleave synaptobrevin to inhibit neuronal transmission incapacitating the host[32]. Shiga and shiga-like toxins from *E. coli* enter cells of the colon and halt protein synthesis leading to cell death and dysentery[33, 34]. The mammalian immune system has also evolved specialized cells (antigen presenting cells, APCs) that internalize extracellular protein into the cytoplasm[35]. Cytoplasmic proteolysis pathways generate peptides that are loaded onto peptide binding proteins (the major histocompatability complex, MHC) which are then displayed on the cell surface for interrogation by T cells[36, 37, 38]. A better understanding of the pathway by which APCs internalize extracellular protein could very likely be exploited for the purposes of vaccine development[39, 40], or engaging the immune system to help fight tumors[41].

I've taken both a top-down and bottom-up approach in seeking to better define the design principles the enable the cytoplasmic access of small miniature proteins. First, using a strategy developed in the Schepartz Lab for the monitoring of protein-protein interactions, I examine the requirements for peptide-MHC association. Next, using a small library of cationic miniature proteins, I probe the influence of positive charge number and orientation on the ability the miniature protein to access the cytoplasm. Using a

novel assay, I identify a cationic miniature protein, called 5.3, carrying a discrete arginine motif that efficiently targets miniature proteins to the cytoplasm. Database searches find that the precise motif identified (an arginine present in positions i, $i+4$, $i+7$, $i+10$, and $i+11$ of an α-helix) is not present in nature, but that similar motifs are present in natural proteins that interact with cellular membranes. Finally, I examine the cellular pathway by which 5.3 reaches the cytoplasm. I find that this miniature protein enters the cell via a dynamin and cholesterol dependent endocytic mechanism and is delivered to Rab5+ early endosomes. In contrast to the shiga-like toxins, and many non-enveloped viruses (which escape to the cytoplasm from the endoplasmic reticulum) and other peptides previously identified as 'cell penetrating', only 5.3 escapes from early endosomes. These findings should enable the future dissection of the precise molecular events underlying cytoplasmic access of peptides and proteins, and may illuminate principles for the engineering of peptides and peptidomimetics that access cytoplasmic targets.

Chapter 2

Use of Bipartite Tetracysteine Display to Probe the Association of Cross-Presented Peptides with the Major Histocompatability Complex

2.1 Background

The mechanistic understanding of biological processes increasingly requires atomic level resolution of changes in protein conformation or association state. Recent results suggest that cells localize signaling processes on discrete membrane[42] or endosomal[43] regions creating a diversity of protein conformations segregated in space. In addition it has long been known that taking protein systems out of their natural context can result in experimental artifacts. Therefore researchers require the ability to interrogate changes in proteins within a cellular environment using a method that provides resolution within time and space. Very few, if any, existing methods can satisfy these demands, and thus new techniques for the interrogation of protein-protein, or protein-peptide interactions

within living cells are required.

Antigen presentation is the process which integrates the responses of innate and adaptive immune cells to generate an immune response, and provides a case-in-point for the above requirements. The process takes place across cellular contacts between specialized antigen presenting cells (APCs) and T cells through the interaction of the major histocompatibility complex (MHC) glycoprotein on the APC and the T cell receptor on the T lymphocytes. While the presence of cytokines in the surrounding media as well as the expression of surface 'costimulatory' proteins on the APC can influence the vigor with which the T cells responds, the primary factor that determines activation of the activation of cytotoxic T lymphocytes (CTL) is the tight association of the T cell receptor with a complex composed of the class I MHC molecule (MHC-1) and an antigenic peptide. Antigenic peptides are derived from proteolytic fragments of proteins both expressed in the cytoplasm (endogenous) and internalized (exogenous). Endogenous antigens are complexed with MHC-1 through specialized chaperones in the endoplasmic reticulum, but the mechanism by which exogenous peptide antigens are loaded onto MHC-1 remains unknown.

2.1.1 The spatial requirements of antigen presentation.

All nucleated mammalian cells utilize the 'classical' MHC-I loading pathway to present endogenous antigen for CTL surveillance. Proteins targeted for degradation by polyubiquitination are cleaved by the proteasome into oligopeptides that are transported into the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP). In the ER, chaperones mediate association of peptide with MHC-1 and the resulting peptide-MHC-1 complexes are transported to the cell surface for display [38].

The presentation of exogenous antigen by MHC-I on the APC surface (a process called cross-presentation [44, 45, 46]) presents several conundrums. Perhaps the most significant is whether exogenous protein antigen must access the cytoplasm for protea-

somal degradation. If exogenous proteins do reach the cytoplasm of APCs, presentation of derived peptide antigens could proceed in the usual fashion. In the absence of specific import machinery, cytoplasmic access of exogenous proteins would represent a severe exception to the permeability barrier of the cell.

At least five distinguishable pathways have been proposed to explain the mechanism by which MHC-1 molecules acquire exogenous peptide antigen[47, 36]. The simplest (A) involves digestion of endocytosed protein in peripheral vesicles that then fuse with recycling endosomes [48] or more acidic compartments (such as the MHC class II loading compartment [49]) that catalyze the exchange of antigenic peptide loading on MHC-1 before transport back to the cell surface. Alternatively, peptides [50] delivered to the cytosol, either by (B) endosomal escape or (C) gap junctions[51] could be imported into the ER by TAP and loaded via the 'classical' MHC-1 pathway. The two remaining pathways involve delivery of intact protein to the cytosol via Sec61. Phagosome-ER fusion (D) may deliver extracellular proteins to the ER; the proteins are then retro-translocated to the cytosol for proteosomal degradation[35]. A similar process (E) could occur in peripheral endosomes that contain Sec61 and TAP, but are distinct from the endoplasmic reticulum[52].

The challenge in understanding the formation of exogenous peptide-MHC complexes is made more difficult by the variety of complexes present, the need for accurate subcellular localization and the small size (but critical nature) of the peptide in comparison to the MHC molecule. In addition, it is the combination of peptide and MHC that must be recognized, because the peptides in isolation are inert, and because APCs continually express, internalize, and degrade multiple alleles of MHC, with some MHC molecules being reloaded and recycled to the cell surface.

Identification of the compartment in which cross-presented peptide MHC complexes first form would provide a way to distinguish experimentally between these hypotheses: pathways B, C, and D predict that cross-presented MHC-peptide complexes should assemble in the ER, whereas pathways A and E do not. Pathways A and B are proteosome-

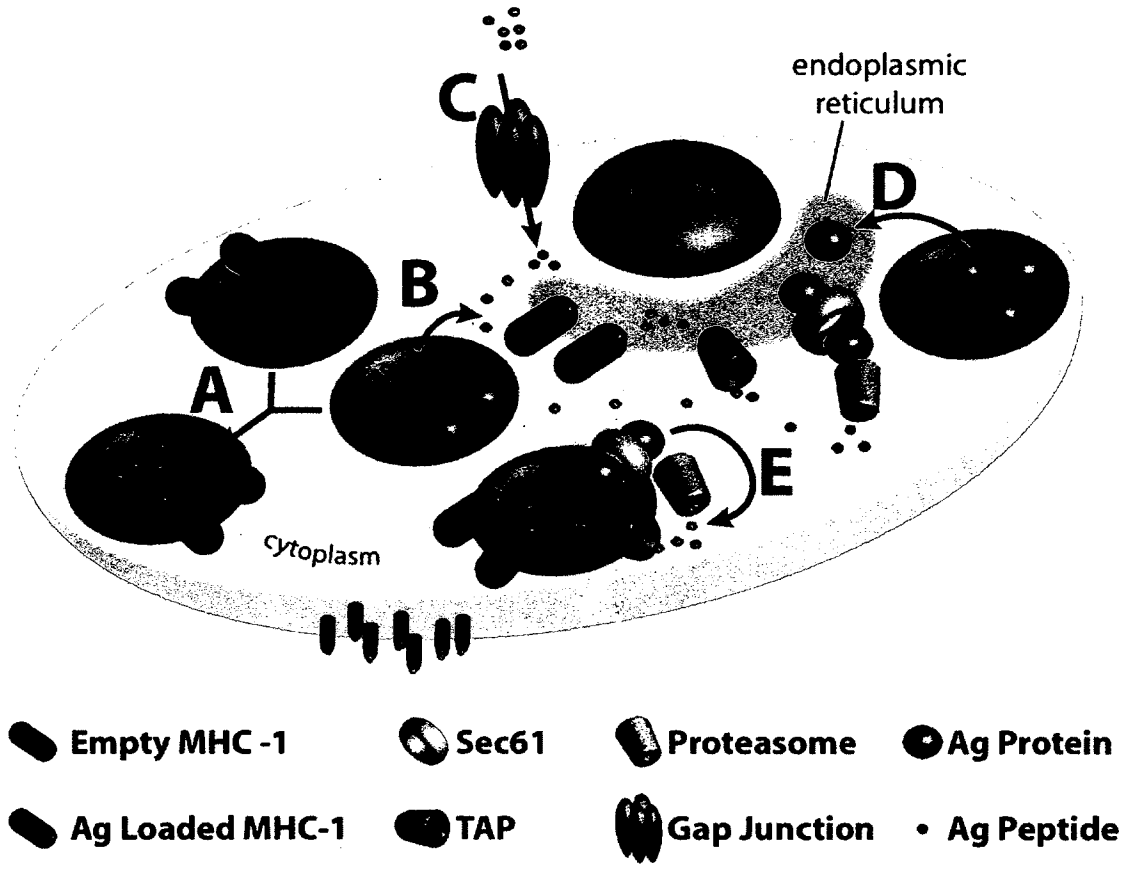


Figure 2.1: Pathways of Cross-presentation. (**a**) Endocytosed protein is degraded in peripheral vesicles creating peptides that encounter empty MHC molecules. The exogenous antigen does not enter the cytoplasm. (**b**) Peptides are delivered to endosomes and then to the cytosol. The mechanism of peptide escape from endosomes has not been determined. (**c**) Peptides from nearby cells are delivered to the cytoplasm directly via gap junctions. (**d**) Phagosome-ER fusion delivers endocytosed protein to the ER. Protein is retrotranslocated to the cytoplasm and is degraded by the proteasome. (**e**) Phagosomes retrotranslocate protein to the cytoplasm, and re-internalize proteasomal peptides for MHC-1 assembly. No communication with the endoplasmic reticulum is required.

independent, and should not be inhibited after addition of the proteasome inhibitor bortezamib. Only model A is TAP-independent. Finally, only pathway D predicts that intact protein antigen should exist in the ER. While experimental evidence for each of these pathways has been observed in various cell line and mouse models, the direct (or indirect) visualization of peptide-MHC complexes in the ER or in the peripheral endomembrane system has not been achieved. Knowledge of dominant pathways could guide development of drugs specifically targeted to blocking presentation of autoantigens, without affecting the rest of the immune system.

2.1.2 Existing approaches to probe protein-protein interactions and their application to cross presentation.

Attempts at antibody-based immunolocalization of cross-presented peptide-MHC complexes have failed[53, 54, 55, 56]. An alternative common approach to studying the subcellular localization of molecules is to tag them with fluorophores synthetically. New small organic fluorophores that (with specialized microscopes) provide sub-diffraction (i.e. 'super'resolution) are available[57]. Covalent attachment of these dyes could provide information on the proximity of peptides derived from exogenous antigen to MHC-1 but even super-resolution experiments are limited to ~20 nm[58], which may define a peptide and MHC molecule being localized to the same membrane region, but may leave ambiguous whether they are complexed together.

Förster resonance energy transfer (FRET) operates on smaller distance scales (1-10 nm) but requires tagging both interacting partners with fluorescent chromophores [59]. Antigenic peptides are small (8 - 11 amino acids), and attaching a fluorescent protein (GFP is 238 amino acids) to MHC near near its binding site may perturb the loading process; placing a fluorescent protein on the C-terminal tail of MHC does not perturb function, but locates the fluorescent protein far from the peptide preventing the possibility of FRET. New methods for detecting protein-protein interactions that rely of the

physical association of proteins (protein complementation assays, see [60]) are being developed, but these require the cooperative folding of protein and peptide to form single domains *in trans*.

Tagging only the peptide with smaller fluorophores, such as fluorescein will not lead to specific detection of MHC-peptide complexes; all of the peptide within a cell would be fluorescent. Small solvatochromatic fluorophores have been incorporated as side chains of unnatural amino acids within peptide MHC II ligands[61]. These molecules show fluorescence enhancement when buried in the MHC-peptide binding groove, but the specificity of the signal in this case was not demonstrated. In addition these molecules showed a fluorescence enhancement roughly two-fold above background, a level not high enough to permit microscopic examination of the cells.

Previous attempts to visualize the formation and trafificking of cross-presented peptide-MHC complexes have been thwarted by the inability of fluorescently labeled antibodies to detect specific complexes within live cells [55]. Bipartite tetracysteine display possesses unique advantages in this regard, as (1) it requires only a 4 amino acid tag, (2) binding is rapid, allowing pulse chase experiments, (3) the complex formed with FlAsH/ReAsH possesses high thermodynamic and kinetic stability[62], and ReAsH (4) can polymerize diaminobenzidine, enabling detection by electron microscopy (EM)[63, 64, 65]. Applying bipartite tetacysteine display to the detection and localization of peptide-MHC complexes should reveal the cellular localization of cross-presented peptide-MHC complexes with high resolution.

2.2 Results

The laboratory of Roger Tsien has previously shown that recombinant proteins containing the linear tetracysteine sequence CCPGCC can be selectively labeled with the cell permeable 'pro-fluorescent' molecules FlAsH (4,5-bis(1,3,2-dithiarsolan-2-yl)fluorescein) and ReAsH (4,5-bis(1,3,2-dithiarsolan-2-yl)resorufin)[66, 63, 64, 67]. FlAsH and ReAsH

are not fluorescent on their own, but become so when bound to peptides or proteins containing a tetracysteine motif[66]. In developing a strategy called bipartite tetracysteine display, the Schepartz Lab split the tetracysteine tag between two proteins so that a tetracysteine binding site is reconstituted upon protein protein interaction, but not when the two proteins remain in isolation. Treating protein complexes with FlAsH or ReAsH resulted in fluorescence increases, but no increase was observed when mutations were introduced that prevented protein-protein interaction [62]. This strategy provides specific advantages in cellular imaging. First the molecular recognition of the tetracysteine binding site by ReAsH requires the cysteine residues to be 5-8 Å apart (see below), delivering spatial resolution that is unachievable by other means in cells. It also does not require the genetic fusion of large fluorescent proteins, and removes the permeabilization step necessary in immunofluorescence[66, 63, 64]. To extend this approach to peptide-MHC-1 recognition, I began by generating a series of peptides and matched MHC-1 variants that contain judiciously placed dicysteine pairs. I then tested the ability of these MHC-1 variants to traffic to the cell surface and to bind to cysteine containing peptides.

2.2.1 A bipartite tetracysteine display strategy to selectively label peptide-MHC-1 complexes.

The original report of the use of bisarsenical profluorescent molecules to label recombinant proteins assumed a target peptide CCRECC was helical[67]. Modeling a well structured α-helix suggested that the thiol-arsenic bonds would be placed within the proximity required to form a tetradentate binding site with the bisarsenical fluorophore. Subsequent optimization of the ligand peptide sequence[68], and the determination of the structure of the optimized peptide-ReAsH complex[69] by NMR reveals that the bisarsenical fluorophore prefers a hairpin structure (Fig. 2.2a). Analysis of the sulfur-sulfur bond distances, as well as the distances between β-carbons is shown (Fig 2.2b). The intra-atomic distances range from ~4-7 Å for the sulfur atoms, and ~5-8 Å for the peptide

β-carbons. This analysis confirms earlier hypotheses[70, 62, 63, 64, 67] that the formation of a tetracysteine binding site, whether linear or bipartite, requires sub-nanometer approximation of the four cysteine residues.

The first structure reported of the human leukocyte antigen (HLA) was of the A2 allele[71, 72], which has served as a model for MHC-peptide loading. In silico mutagenesis of these complexes to introduce potential FlAsH/ReAsH binding sites suggests two positions – 179 and 183 on HLA-A2 – as ideal for the requisite Cys-Cys pair: the α-carbons of these side chains are located approximately 6 Å from each other and 6 Å from the two cysteines in the peptide YLLCCNYKL, which binds HLA-A2 with high affinity ($K_D = 2\,\mathrm{nM}$) [73, 74]. The α-carbon of other amino acids located within the α-helices that border the peptide binding groove were also closely approximated to the α-carbons on the bound antigenic peptide. These include R89, K90, A93, V97, T100, K170, A174, and A182. We searched online databases[75] for other cysteine pair-containing peptides that bind HLA-A2 with high affinity. Identified ligands include IIITCCLLSV ($K_D = 6.3\,\mathrm{nM}$) and MICCDSRIVV ($K_D = 24.6\,\mathrm{nM}$). Pairing selected HLA-A2 variants with peptides containing C_2 pairs in different locations along the peptide binding groove should allow for examination of the proximity requirements for the formation of a bipartite C_4 motif between the MHC variant and peptide. For example, I predict that K90C/A93C and Y183C/T187C should form a suitable FlAsH/ReAsH binding motif in combination with IIITCCLLSV or MICCDSRIVV, respectively. Additional sites on HLA-A2 suitable for C_2 installation are shown in figure 2.2d. Wild type HLA-A2 modified with an intact tetracysteine (C_4) tag serves as a positive control whose ability to bind FlAsH or ReAsH should be independent of added peptide. Wild-type HLA-A2 lacking a C_2 motif serves as a negative control.

2.2.2 Expression analysis of generated MHC-1 Variants

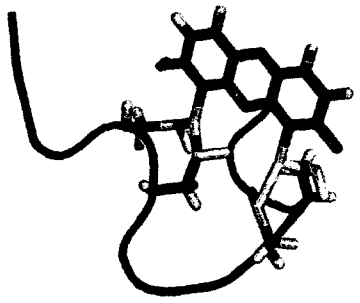
I took note of the fact that the insertion of cysteine residues within the MHC heavy chain may not be innocuous to the function of the protein. During the synthesis of MHC-1, nascent proteins are tightly complexed with chaperones in the ER[77, 78]. Release from the ER requires proper folding (for which peptide loading is a prerequisite[79]), and so the presence of MHC-1 on the surface of cells implies that these molecules are folded properly and loaded with peptide[80, 81]. Wild-type HLA-A2, when tagged with GFP and expressed in cultured epithelial cells, is present both at the surface and within intracellular vesicles[49]. To determine whether the introduction of cysteines near the peptide binding groove induced perturbations to the subcellular localization of MHC-1, we expressed GFP fusions of dicysteine containing HLA-A2 variants in HeLa cells and examined the cells by microscopy (Fig. 2.3).

HeLa cells were transfected with wild-type HLA-A2 tagged to GFP (HLA-A2-GFP) or variants containing engineered dicysteines. The presence of expressed HLA-A2 was determined by the observation of GFP fluorescence. To determine whether the transfected HLA-A2 variants were present on the cell surface, transfected cells were labeled with phycoerythrin (PE) conjugated BB7.2, a monoclonal antibody specific for the folded conformation of HLA-A2. As can be seen in figure 2.3b, HeLa cells transfected with wild-

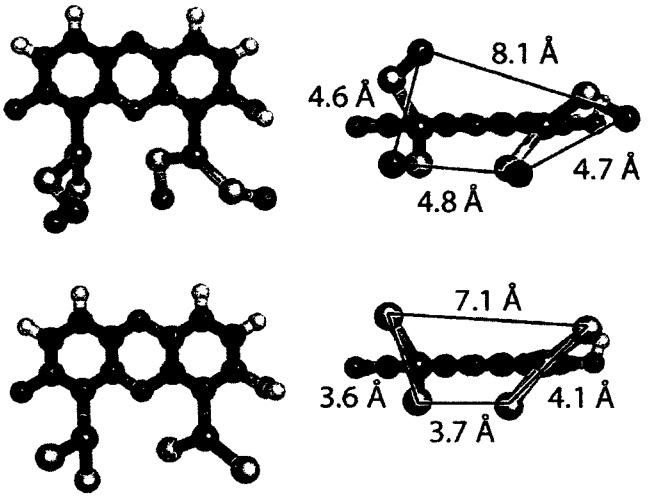
---

Figure 2.2 *(following page)*: A bipartite tetracysteine-based strategy to detect peptide-MHC complexes. (**a**) The structure of ReAsH bound to the peptide sequence FLNC-CPGCCMEP, as determined by NMR (personal communication from RY Tsien). (**b**) The ReAsH molecule from **a**, shown without the peptide, but with bound thiol and β-carbon (above) or only bound thiol (below) ligands. The distances between β-carbons and between sulfur atoms is shown. (**c**) The structure of the peptide binding groove of HLA-A2[76] as determined by X-ray crystallography (PDB structure 1JHT). The sites of cysteine incorporation are shown. (**d**) The combination of variants made. Each variant contains two engineered cysteine residues. (**e**) Possible orientations of tetracysteine motifs generated upon peptide binding. Each possible tetracysteine site is shown in a different color. None of the distances shown are > 10 Å from α-carbon to α- carbon. (**f**) A close up view of tetracysteine site 1 (shown in **e**). The inter-β-carbon distances are shown. This view is rotated 180° from the view in e in order to better illustrate the close apposition of the peptide with the α-helix that forms the peptide binding grove.

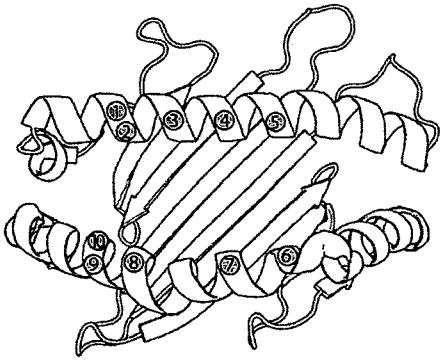
a



b



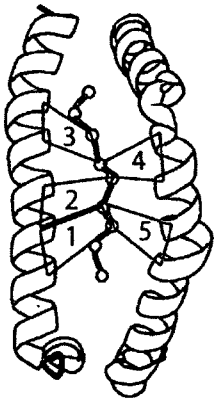
c



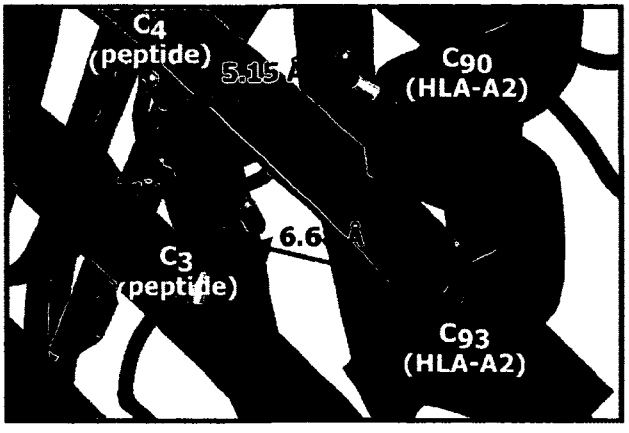
d

Position #	Residue #	Variants Made					
1	R89	+	+				
2	K90	+					
3	A93		+				
4	V97			+			
5	T100			+			
6	K170				+		
7	A174				+		
8	Q179					+	+
9	A182					+	
10	Y183						+

e



f



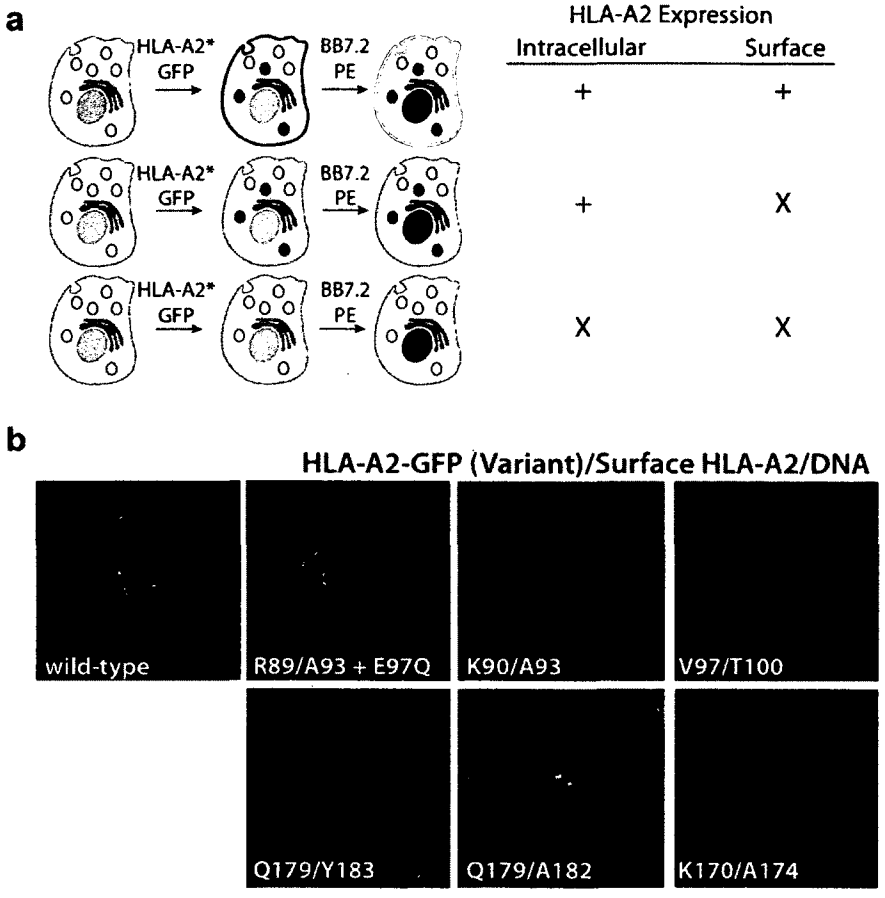


Figure 2.3: Epifluorescence microscopy showing the subcellular distribution of GFP-tagged dicysteine containing HLA-A2 variants. (**a**) Experimental Schema. HeLa cells are transfected with wild-type HLA-A2 tagged to GFP (HLA-A2-GFP) or HLA-A2 variants containing engineered dicysteines. Cells are then stained with BB7.2-PE, an antibody that recognizes folded HLA-A2[82], without permeabilization. The presence of red fluorescence on the cell surface indicates the presence of properly folded HLA-A2, which when overlapped with green can appear yellow. Variants of HLA-A2-GFP that express but do not fold, may be retained in the ER or golgi, and should not lead to red surface staining with BB7.2-PE. Some variants of HLA-A2-GFP may be unstable and rapidly degraded. In this case, neither green nor red fluorescence will be observed. (**b**) Cell micrographs prepared as described in **a**. Wild-type HLA-A2-GFP is expressed intracellularly and on the cell surface after transfection in HeLa cells. R89/A93, K90/A93, Q179/A182, and K170/A174 variants also express on the cell surface. Surface expression of Q179/Y183 is low. The expression of the V79/T100 variant is undetectable.

type HLA-A2-GFP display both green intracellular fluorescence and red fluorescence on the cell surface. This is indicative of the proper folding and surface expression of the transfected HLA-A2 allele. Cells transfected with A2 variants containing engineered dicysteines at positions R89/A93 or Q179/A182 show patterns of intracellular and surface expression that are qualitatively similar to wild-type. HeLa cells transfected with an A2 variants containing engineered dicysteines at either positions K170/A174 or Q179/Y183 show both intracellular GFP and surface BB7.2 labeling, but increased green fluorescence signal is located within the intracellular reticular network, indicating the possibility of inefficient folding. The A2 variant containing dicysteines at position V79/T100 shows no GFP fluorescence or BB7.2-PE labeling, indicating that this variant is not expressed.

To quantitatively assess whether A2 variants showing surface expression by microscopy are present at levels similar to wild-type, HeLa cells transfected with wild-type HLA-A2-GFP or dicysteine containing variants were labeled with BB7.2-PE and examined by flow cytometry (fig 2.4). Living cells were identified by their characteristic pattern of forward scatter (FSC) and side scatter (SSC), and their ability to exclude propidium iodide. GFP expressing cells were identified by the presence of green fluorescence and their levels of surface BB7.2 staining were compared. As can be seen in figure 2.4b, HeLa cells transfected with wild-type HLA-A2-GFP show high levels of surface BB2.7 staining (green line), but untransfected cells (black line) and cells not exposed to the antibody (grey shaded histogram) do not, confirming that HeLa cells are HLA-A2-negative[83] in the absence of transfection, and that transient transfection induces surface A2 expression. Comparing the histograms shown in figure 2.4c-g to the histogram in figure 2.4b shows that HeLa cells transfected with variants containing engineered dicysteines at sites R89/A93, K90/A93, K170/A174 and Q179/A182 express surface levels of HLA-A2 that are similar to levels obtained for the wild-type allele. HeLa cells transfected with HLA-A2 containing an engineered dicysteines at position Q179/Y183 show either reduced surface expression or reduced antibody binding. (The latter is a possibility, because the BB7.2 epitope is composed of the C-terminus of the α2-

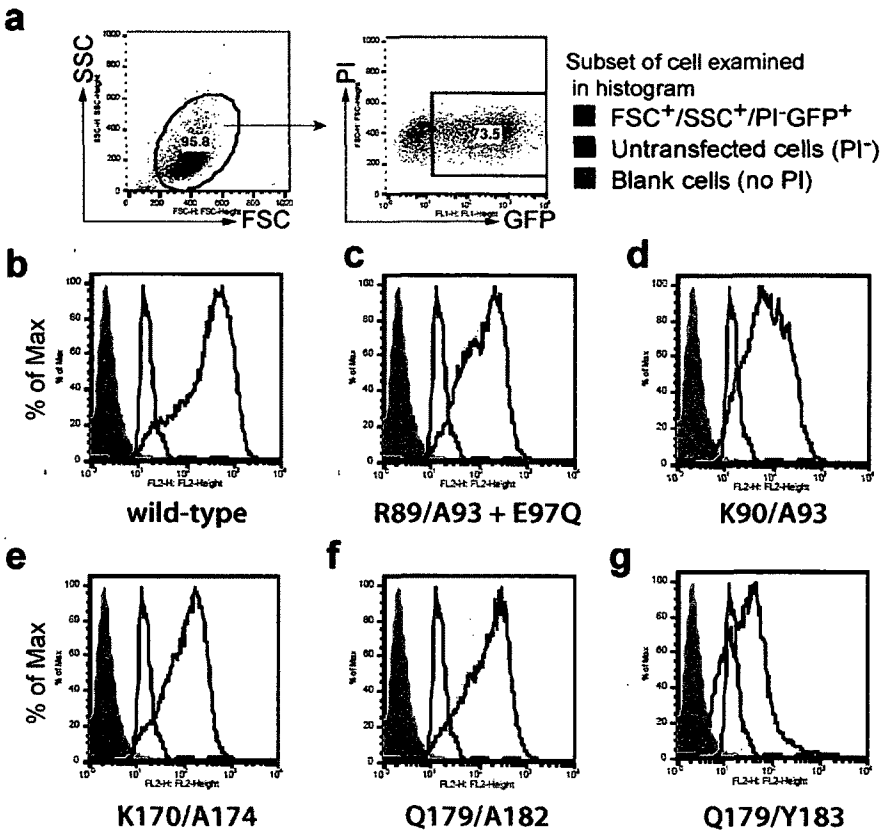


Figure 2.4: Surface expression of dicysteine containing HLA-A2 variants as assessed by flow cytometry. HeLa cells transfected with HLA-A2-GFP or dicysteine containing variants were stained with BB7.2-PE as shown in figure 2.3 before analysis by flow. (a) Process of analysis. Cells were gated on forward scatter (FSC) and side scatter (SSC) to define a 'live cell' population. Cells from the FSC/SSC gate were examined for GFP expression and propidium iodide uptake. Cells there were PI^- (live) and GFP positive were selected for inclusion in the histograms shown in **b**. (**b-g**) The cell fluorescence due to BB7.2-PE staining of GFP+ cells (green line) is compared to cells transfected with GFP and treated with PI (PI^-), black line, as well as untransfected cells, not treated with PI (gray shaded histogram).

helix, wherein Q179 and Y183 are located, and portions of its underlying β-strand[84].) This analysis confirms the conclusions following analysis by epifluorescence microscopy, namely, that the introduction of dicysteines at positions R89/A93, K90/A93, K170/A174 and Q179/A182 do not severely perturb the expression, folding, or transport of HLA-A2 molecules. The same may also be true for the A2 variant with the dicysteine introduced at position Q179/Y183.

2.2.3 Recognition of tetracysteine motifs within oxidizing cellular environments.

The binding of ReAsH or FlAsH to peptide tetracysteine motifs requires arsenic-thiol exchange. Because FlAsH and ReAsH will not react with oxidized disulfides[85], the tetracysteine must be completely reduced. Were the above dicysteine containing HLA-A2 variants to assemble with dicysteine containing peptides, the reconstituted tetracysteine motif would be present within the lumen of the endoplasmic reticulum, an oxidizing environment that promotes disulfide bond formation[86]. The addition of small thioalkanes can reduce intralumenal thiols[87] revealing tetracysteine motifs for FlAsH and ReAsH labeling[64]. In fact, careful choice of reductant can allow for pulse chase experiments, distinguishing tetracysteine motif-containing proteins that are sensitive to cell permeant versus cell impermeant reductants, or those that are located within reducing vs oxidizing environments[64].

To determine whether I could detect ER lumenal tetracysteine containing proteins, I transfected HeLa cells with an expression vector encoding the HLA-A2 amino acid sequence with the peptide sequence CCPGCC prepended to the N-terminus (C4-HLA-A2-GFP). This C4-HLA-A2 should display labeling with ReAsH that is independent on the formation of a bipartite tetracysteine motif composed of MHC and peptide. To confirm that treatment with reducing agents could lead to the successful labeling of a tetracysteine motif present within the lumen of endosomes, these cells were treated with

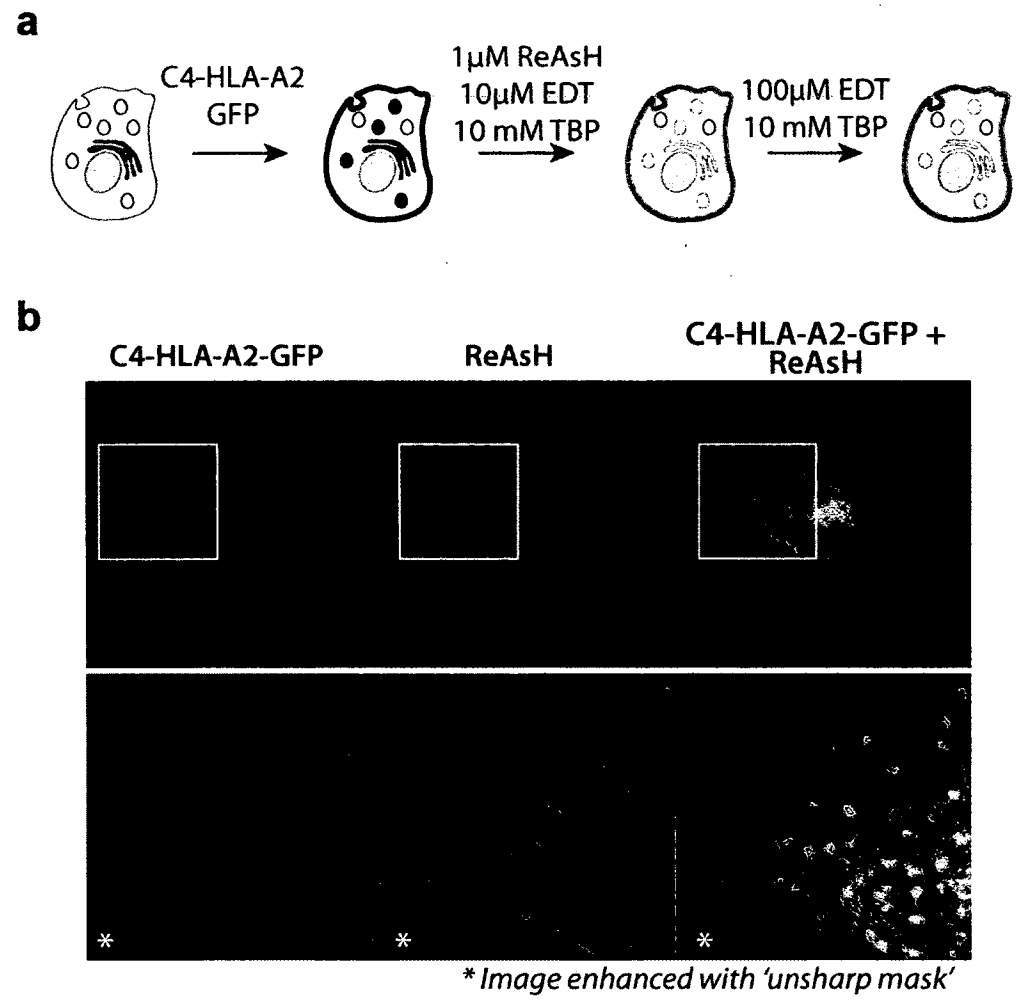


Figure 2.5: Detection of tetracysteine motifs within oxidizing environments. HLA-A2 was tagged on the N-terminus (within the ER lumen) with the peptide sequence CCPGCC. (**a**) Labeling schema. HeLa cells were transfected with C4-HLA-A2-GFP for 24 hr before treatment with 1 μM ReAsH, 10 μM EDT, and 10 mM TBP for 20 min. Cells were then washed with HBSS and incubated for 20 min with 100 μM EDT and 10 mM TBP to reduce non-specific ReAsH labeling before imaging by epifluorescence microscopy. (**b**) Images of a HeLa cells transfected with C4-HLA-A2-GFP and labeled with ReAsH as described in **a**. Fluorescence due to GFP is shown in green, while fluorescence due to ReAsH is shown in red. An overlay of the two images is shown on the right, where overlap of the green and red signals appears yellow. In the periphery of the cell, the red ReAsH signal is concentrated in regions of GFP fluorescence. A zoomed image of the region within the white box is shown below.

ReAsH and 10 mM tributylphosphine (TBP), a cell permeable reducing agent previously shown to reduce thiols present within the golgi lumen[64]. In addition, 10 μM ethanedithiol (EDT) was included in the labeling mixture as a competitor for the non-specific interaction of ReAsH with cellular protein thiols. After a washing step with a higher concentration of EDT (100 μM), red ReAsH signal was visible within endosomes. A diffuse background of red signal within the cell body was also present. Examination of the GFP fluorescence showed a pattern of bright puncta combined with thin branching and anastomosing tubules forming a reticular network concentrated near the nucleus. This pattern is similar to that observed after transfection with wild-type HLA-A2-GFP, although with decreased signal originating from the plasma membrane (compare figure 2.3 with figure 2.5). When the images of GFP fluorescence and ReAsH signal were overlaid, the areas of more intense ReAsH signal colocalized with GFP positive vesicles, confirming that the presence of ReAsH signal from these vesicles was due to the presence of the C4-tagged HLA-A2 molecules.

2.2.4 Dicysteine HLA-A2 variants are not loaded with dicysteine-containing A2-binding peptides.

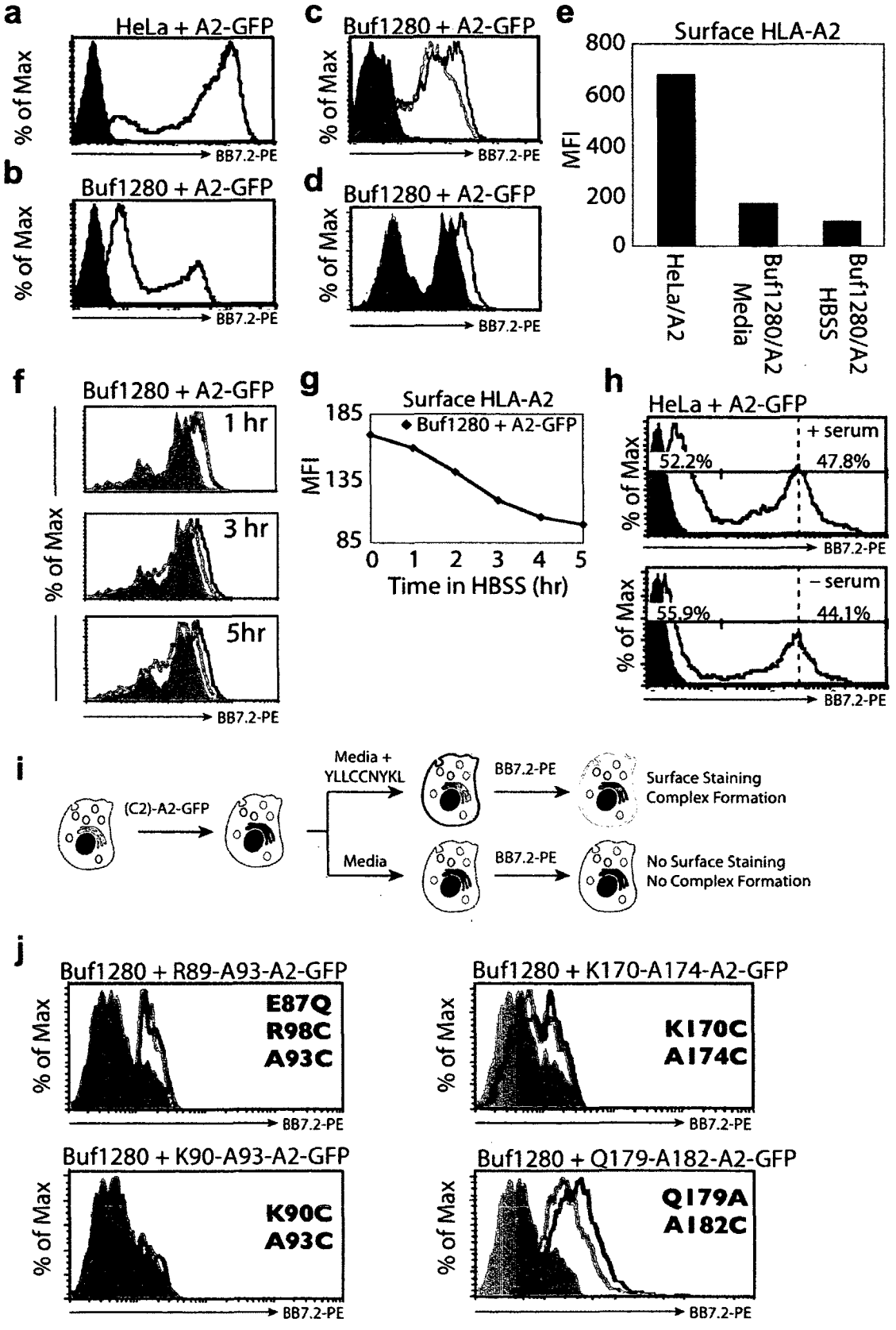
Initial trials of 'feeding' dicysteine containing peptide to HeLa cells transfected with dicysteine containing HLA-A2 molecules led to high cellular background, and no appreciable signal. To bias the system toward the formation of peptide-MHC complexes containing bipartite tetracysteine motifs, we relied on observations that in cells harboring loss-of-function mutations in the transporter associated with antigen presentation (TAP), expressed MHC molecules are retained in the endoplasmic reticulum, but the addition of exogenous peptides can lead to peptide loading and stabilization of MHC surface expression[88, 74]. Three such TAP^{-} cell lines were available as kind gifts of the Cresswell laboratory, RMA-S (mouse H-2^b), T2 (human HLA-$A2^{+/+}$), and Buf1280 (human, HLA-$A2^{-/-}$, [89]). Both RMA-S and T2 are lymphocytic in origin, and thus are

challenging targets in experiments using microscopy for the determination of subcellular localization. Buf1280, on the other hand, is a melanoma cell line (epithelial in origin) that adopts a broad flat morphology preferable for microscopy. This cell line harbors a 2 bp insertion mutation that inactivates HLA-A2 expression, but transfection with wild-type HLA-A2 and TAP rescues the surface expression defect[89]. Thus we chose the Buf1280 cell line to test whether the dicysteine containing peptides identified in database searches could bind to the dicysteine containing HLA-A2 variants.

To begin, we first used flow cytometry to confirm that expression of HLA-A2 in Buf1280 cells led to low levels of surface expression and was unstable. HeLa cells transfected with HLA-A2-GFP show two histogram peaks when stained with BB7.2-PE. A high staining peak (characteristic of transfected cells) and a low staining peak (characteristic of untransfected cells). When compared to the levels of surface expression of wild-

---

Figure 2.6 *(following page)*: Surface stabilization of dicysteine containing HLA-A2 variants in cells lacking TAP. (**a**) HeLa cells transfected with wild-type HLA-A2-GFP express high levels of stable surface HLA-A2 as detected by BB7.2-PE staining. (**b**) Buf1280 cells, when transfected with wild-type HLA-A2-GFP show low levels of surface expression. For panels **a** and **b**, cells showing FSC/SSC profiles and PI negative staining (characteristic of living cells) are included in this histogram. Expression of HLA-A2 on the surface of Buf1280 cells can be stabilized by incubation for 4 hrs at 37°C with media supplemented with 10% serum (**c**, blue histogram, gated GFP+) or10 μM YLLCCNYKL, an HLA-A2 binding peptide (**d**, grey histogram, gated GFP+). Cells incubated in serum free media (red histogram in **c**, blue shaded histogram in **d**, both gated GFP+) and GFP- untransfected cells (grey shaded histograms) are shown as controls. (**e**) Surface stabilization of HLA-A2 by serum on Buf1280 cells compared to the levels of surface expression in HeLa cells. Surface expression of HLA-A2 is unstable. (**f**) After incubation in media supplemented with serum overnight, HLA-A2-GFP expressing Buf1280 cells were incubated in HBSS without serum for varying times. (**g**) Mean fluorescence intensity as a function of time. (**h**) HeLa cells expressing HLA-A2 and incubated with or without serum for 4 hr at 37°C show similar levels of surface HLA-A2. (**i**) A schema for the determination of the surface stabilization of dicysteine containing HLA-A2 variants by dicysteine containing peptides. Buf1280 cells transfected with HLA-A2 variants are incubated overnight in serum free media with or without YLLCCNYKL peptide before staining with BB7.2-PE to measure the levels of surface HLA-A2. (**j**) Measurement of HLA-A2 surface stabilization on Buf1280 cells by dicysteine containing peptide. In every case examined, the levels of surface HLA-A2 were the same (R89/A93, K90/A93, K170/A174) or lower (Q179/A182) in the presence of peptide (red histograms) than in its absence (blue histograms).


a
HeLa + A2-GFP
% of Max
BB7.2-PE
b
Buf1280 + A2-GFP
% of Max
BB7.2-PE
c
Buf1280 + A2-GFP
% of Max
BB7.2-PE
d
Buf1280 + A2-GFP
% of Max
BB7.2-PE
e
Surface HLA-A2
MFI
800
600
400
200
0
HeLa/A2
Buf1280/A2 Media
Buf1280/A2 HBSS
f
Buf1280 + A2-GFP
1 hr
3 hr
5hr
% of Max
BB7.2-PE
g
Surface HLA-A2
Buf1280 + A2-GFP
MFI
185
135
85
0 1 2 3 4 5
Time in HBSS (hr)
h
HeLa + A2-GFP
+ serum
52.2%
47.8%
% of Max
BB7.2-PE
– serum
55.9%
44.1%
% of Max
BB7.2-PE
i
(C2)-A2-GFP
Media + YLLCCNYKL
BB7.2-PE
Surface Staining
Complex Formation
Media
BB7.2-PE
No Surface Staining
No Complex Formation
j
Buf1280 + R89-A93-A2-GFP
E87Q
R98C
A93C
% of Max
BB7.2-PE
Buf1280 + K170-A174-A2-GFP
K170C
A174C
% of Max
BB7.2-PE
Buf1280 + K90-A93-A2-GFP
K90C
A93C
% of Max
BB7.2-PE
Buf1280 + Q179-A182-A2-GFP
Q179A
A182C
% of Max
BB7.2-PE

type HLA-A2 when transfected into HeLa cells (fig. 2.6a), the same allele transfected into Buf1280 cells showed lower expression levels (fig. 2.6b). This low level of expression could be increased by incubating the cells in media containing serum, or in serum free media supplemented with 10 μM YLLCCNYKL, a high-affinity HLA-A2 binding peptide (fig. 2.6c and d). While incubation with peptide increases surface expression, the levels are still low compared to the steady state level in HeLa cells (fig. 2.6e). Peptides stabilize levels of surface expression transiently; removing the cells from media containing serum results in a time dependent decrease in surface folded HLA-A2 (fig. 2.6f), with a half-life of approximately 2.5 h (fig. 2.6g). By contrast, when HeLa cells expressing HLA-A2 are incubated in the presence or absence of serum for 4 hr, the levels of surface HLA-A2 are similar (fig. 2.6h). These results confirm that expression of HLA-A2 on the surface of $TAP^{-/-}$ cells is not long lived, but can be increased through the addition of exogenous peptide, either in the form of serum, or in the form of specific HLA-A2 binding peptides, such as YLLCCNYKL.

To test whether YLLCCNYKL stabilized dicysteine containing variants of HLA-A2, I transfected Buf1280 cells with plasmids encoding HLA-A2 incorporating dicysteine residues at positions R89/A93, K90/A93, K170/A174 or Q179/A182. The cells were then incubated in serum free media or in serum free media supplemented with 10 μM YLLCCNYKL. After overnight incubation at 37°C, the levels of surface HLA-A2 were examined by flow cytometry (for schema see fig. 2.6i). As can be seen in figure 2.6j, the levels of surface expression of HLA-A2 variants incorporating dicysteines at positions R89/A93, K90/A93 or K170/A174 were similar regardless of the presence or absence of the HLA-A2 binding peptide. For the HLA-A2 variant harboring cysteine residues at positions Q179/A182, the addition of the peptide decreased surface expression. Taken together, these results indicate that while the YLLCCNYKL peptide stabilizes the surface display of wild-type HLA-A2, it fails to stabilize the surface display of any of the 2C variants tested.

The absence of surface stabilization of HLA-A2 when Buf1280 cells were treated with

2C containing peptide suggest that the peptide failed to form complexes with the HLA-A2 heavy chain. These results make it unlikely that a bipartite tetracysteine motif is formed by the close association of heavy chain and exogenously applied peptide, and make it unlikely that the specific staining will be observed after treating these cells with ReAsH. Indeed, in other experiments, the addition of dicysteine containing peptides substantially increased the cellular background after ReAsH staining, making the likelihood of observing specific peptide-MHC complexes using this method very small.

2.3 Discussion

Dicysteine containing HLA-A2 side chains were expressed in $TAP^{-/-}$ cells, but their surface levels were very low relative to wild-type HLA-A2, even though the surface expression level of these alleles were comparable in $TAP^{+/+}$ (HeLa) cells. These data are in line with findings that TAP plays a role in the folding of MHC-1 heavy chains[38], not just peptide loading, and supports the hypothesis that the addition of cysteines to the heavy chain represents a challenge to the folding of MHC-1 that is made more difficult by the absence of TAP. The MHC-1 loading machinery consists of the MHC-1 heavy chain, β_2-microglobulin, TAP, tapasin, ERp57, and the chaperones and calreticulin. Exclusion of ERp57, an ER resident protein disulfide isomerase, from the MHC-1 loading complex through the expression of a tapasin mutant[90] or ERp57 deletion[78], results in aggregate formation. These aggregates contain tapasin, unfolded MHC-1, as well as other proteins. One current model is that tapasin forms a bridge connecting TAP and MHC-1 with ERp57 and calreticulin[91]. In this model, ERp57, MHC-1 and tapasin should associate independently of TAP. In fact, *in vitro* experiments have shown that tapasin and ERp57 are the minimal functional unit of the peptide loading complex[77]. Thus, why the deficiency of TAP should lead to a selective decrease in the surface expression of dicysteine containing HLA-A2 molecules, even in the presence of exogenously added HLA-A2 binding peptides, remains a mystery.

When DNA encoding dicysteine containing HLA-A2 molecules was transfected into $TAP^{+/+}$ HeLa cells, the proteins were detected on the cell surface, but whether these molecules were loaded with peptide was not tested. I found that the addition dicysteine containining HLA-A2 binding peptides did not stabilize dicysteine containing HLA-A2 heavy chains on the surface of $TAP^{-/-}$ cells, even though these same sequences did stabilize the surface expression of wild-type HLA-A2. One possibility for this difference is that the combination of mutations within dicysteine containing HLA-A2 heavy chains and dicysteine containing HLA-A2 binding peptides prevent association and stable folding of the complex.

Alternatively, these data may point to an additional role for TAP in the MHC-1 peptide loading complex. Unfolded MHC-1 binds to calreticulin and tapasin for incorporation into the loading complex[91]. Even though MHC-1 does not bind to TAP in the absence of tapasin[91], it remains unstudied whether efficient recruitment of MHC-1 into a complex with ERp57 and tapasin is affected by the absence of TAP.

The hypothesis that MHC-1 recruitment into the peptide loading complex is slower in the absence of TAP would explain the findings above. In the case of the wild-type allele, such efficient recruitment may not be necessary, because the protein may fold with an efficiency high enough to maintain reasonable expression levels. On the other hand, in the case where cysteine residues are added to the heavy chain, the absence of swift recruitment to the protein disulfide isomerase ERp57 may lead to a kinetic trap of improperly folded HLA-A2. Rapid recruitment to the tapasin-ERp57 loading machinery in the presence of TAP may occur before such misfolding occurs, in the absence of TAP, the expression of HLA-A2 with additional cysteines thus may be severely hindered.

While this study was ultimately unsuccessful in defining the MHC-1 cross presentation pathway, it did provide several important lessons. These lessons take two forms: principles for choosing target pathways likely to be successfully interrogated by bipartite tetracysteine display, and required features of strategies interrogating cross presentation.

2.3.1 Future Directions I: Strategies for interrogating cross presentation

As illustrated earlier, cross presentation represents a number of particularly vexing challeges. The activation of T cells may occur with very few peptide-MHC complexes (≤ 3 for CD8+ T cell killing [92, 93]), so the detection must be exquisitely sensitive. These peptide-MHC complexes may be distributed across multiple areas of the cell, requireing the ability to detect single molecules. Indeed it is now appreciated that T cells could recognize single molecules long before microscopes could do so[94]. Thus most of the productive approaches to detecting the cellular pathways have used T cell responses as a functional read-out of whether pathways for cross-presentation remain intact after experimental purterbation. And specific theories regarding the cellular trafficking of membrane and endosomal contents are tested as they arise.

Three studies[95, 96, 97] suggested that the fusion of the endoplasmic reticulum (ER) with maturing phagosomes could explain the cross presentation of antigens derived from intact protein. The theory, now wide-spread, is that ER retrotranslocation machinery would pull denatured proteins from the ER to the cytoplasm. A functional role for the ER associated degradation (ERAD) machinery in cross-presentation was provided soon after the initial studies [98, 99], however the notion that the membrane that fused with nascent phagosomes originated from what was canonically considered the ER was challenged[100]. Thus a picture in which the ERAD machinery may be functional outside the ER has emerged, or that phagosomes may fuse with ER-like vesicles containing enzymes thought to be resident only in the ER. Recent studies have suggested that the core component of the ERAD machinery, p97, can be recruited to early endosomes by poly-ubiquitinated mannose receptor[101]. Functional studies arresting the maturation of early or late endosomes may shed light on vesicular pathways required for cross presentation.

Are exogenously derived peptides segregated from endogenous antigen during cross-presentation? At least one group[96] has argued that proteasomes are associated with

the cytoplasmic side of phagosomes and the proteolysis products may be imported back into phagosomes from which they were previously exported. This would provide a way to segregate two classes of MHC-1, one that was loaded in the phagosome, and another that was loaded in the ER. Presumably the antigenic origins of these two populations are distinct, and the cell could regulate their trafficking differently. No data exists to support this model. However, research has shown that some MHC-1 does pass out of the ER into phagosomes where it is loaded and traffics to the cell surface without passing through the golgi (identified through endoglucosidase H sensitivity)[99]. Thus at least two populations of peptide loaded class I proteins exist: those that have passed through the golgi, and those that have not.

All of the pathways enumerated in figure 2.1 are likely to be operable at different times and in different cells. Nevertheless, the mechanism of cross-presentation of antigens derived from endocytosed proteins is perhaps the most interesting from a cell biological standpoint. Despite its importance from the standpoint of vaccine development and understanding basic cellular processes, no available techniques are truly well suited to answer the questions presented in the beginning of this chapter.

With improvements in live cell microscopy, photodynamic/photoactivation approaches may be able to determine which population of MHC heavy chains carry cross presented peptides. In an alternative approach, one could imagine purifying high-mannose heavy chain and examining the mass spectra of the eluted peptides. This may provide one way to determine whether cross-presented antigen is selectively loaded on phagosomal MHC-1 rather than in the ER.

2.3.2 Future Directions II: Targets of choice for bipartite tetracysteine display

The experiments outlined here were targeted not only at gaining a better understanding of antigen cross-presentation, but also at gaining a better understanding of methods and criteria for the use of bipartite tetracysteine display. One principle that has emerged

from *in vitro* protein labeling studies is the need for site flexibility in selecting protein regions in which to incorporate bipartite tetracysteine motifs[102]. The factors presenting challenges during this work have illustrated guidelines for future choices of cellular systems and pathways likely to be successfully interrogated by this technique.

At least four priciples for success have emerged. Systems likely to be successfully interrogated by bipartite tetracysteine display have (1) high levels of expression of a homogenous target, this leads to high concentrations of tetracysteine targets for labeling. (2) It is preferable that biological manipulations can leads to the presence versus the absence of the formation of the bipartite tetracysteine motif, this removes the need to distinguish between 'levels' of labeling, a distinction that requires a dynamic range roughly 5-10 times larger than that required to determine a binary (present vs absent) response[103]. (3) The presence if the tetracysteine motif near the cell surface allows the use of total internal reflection (TIR) fluorescence microscopy (TIR-FM), a technique that selectively excites the ~100 nm of the cell closest to the water-glass interface[104], and can eliminate much the cytoplasmic background seen after ReAsH staining. (4) The tetracysteine motif should be located in the cytoplasm, to remove the need for acute reduction of oxidized thiols.

In retrospect, the bipartite strategy outlined above for the detection of crosspresented MHC-peptide complexes incorporated none of these priciples. With advances in microscopic techniques and approaches, advanced image process, new profluorescent fluorophores, or a better understanding of the 'optimal' bipartite tetracysteine motif, some or all of these restrictions could be relaxed in the future.

2.4 Experimental Methods

2.4.1 *In silico* mutagenesis and structural analysis of MHC-1 and ReAsH-FLNCCPGCCMEP

To determine the intra-atomic distances likely to for a suitable tetracysteine motif for the binding of ReAsH, the PDB file containing the lowest energy conformer from the NMR-determined structure of ReAsH complexed to the peptide FLNCCPGCCMEP was loaded into MacPyMol (`www.pymol.org`). The PDB coordinates were provided by Roger Tsien. The distances between atoms were determined using the 'Measurement' tool in PyMol. The distances between sulfur atoms or β-carbons (cysteine side chains) were measured and are shown in figure 2.2.

To determine possible sites of mutagenesis on the HLA-A2 heavy chain, the PDB file 1JHT [76] was examined. This structure contains HLA-A2 bound to the antigenic peptide ALGIGILT. Surface residues exposed residues (surface exposure $\geq 2.5\AA^2$) were identified using the `findSurfaceResidues` (script available at `http://www.pymolwiki.org/index.php/Category:Script_Library`). In this structure the antigenic residues in IGIL (amino-acids 4-7) are surface exposed. Residue 2 (L) shows a surface exposure of > 2.5 Å, however, the leucine side chain is pointing into the peptide binding groove, while residue 3 (G) shows a surface exposure of < 2.5 Å, but the backbone atoms are unburied. The assignment of this residue as buried may be because the side chain of residue 4 (I) is pointing out from the peptide binding grove, covering this amino acid. Thus residues 3-7 were selected as possible sites for cysteine incorporation. Next, residues containing atoms within 8Å of any atom within peptide residues 3-7 were selected. Surface exposed residues within this selection were chosen as an initial set from which to design possible cysteine incorporation sites. These residues are listed in table 2.1. Modeling the distances from the $\alpha-$carbon residues revealed that some combinations of these side chains existed in preferable geometry, and these were selected for mutagenesis.

Residue	Reason for Inclusion/Exclusion	Could Pair With
R89*	Included	A93
K90	Included	A93
A93	Included	R89, K90
H94	Points away fom binding groove	
V97	Included	T100
R98	Points away fom binding groove	
T100	Included	V97
T146	Too far from peptide	
T147	Too far from peptide	
K170	Included	A174
W171	Side chain buried	
A174	Included	K170
H175	Points away fom binding groove	
V175	Kink in α-helix, no obvious pair	
A177	Included	Q179
Q179	Included	A177, A182
A182	Included	Q179
Y183	Included	Q179

Table 2.1: Residues of HLA-A2 selected for installation of cysteines. Residue R89 was selected by visual inspection, the other residues were selected by the criteria that they contain $\geq$ 2.5 Å of solvent exposed surface, and contained atoms < 8 Å from from atoms contained within the residues 3-7 of the antigenic peptide.

2.4.2 Generation of dicysteine containing HLA-A2-GFP variants

Mutations within the helical segments of HLA-A2 adjacent to the peptide binding groove were installed using QuikChange mutagenesis. Primer design was accomplished using the QuikChange primer design program on the Stratagene website. Primer sequences are listed in table 2.2. A plasmid containing the cDNA for HLA-A*0201 was a kind gift of Ira Mellman and Lelia Delmarre, and was used as a template for the mutagenesis reactions.

To install an N-terminal tetracysteine motif, QuikChange mutagenesis was used to instal an NgoMIV site within the coding region of HLA-A*0201, C-terminal to the signal sequence. The installation of this restriction site did not alter the amino acid sequence of the protein. Next, after desalting using G-25 microspin columns, oligos NC4-Tag and NC4-Tag_R were annealed in T4 DNA ligase buffer. The annealed oligo pair and the plasmid containing the inserted NgoMIV were both digested with the NgoMIV. The plasmid was treated with antarctic phosphatase, and both the insert and plasmid were purified using a PCR cleanup kit from Qiagen. The insert was then ligated into the vector using T4 DNA ligase, and transformed into *e.coli*. Positive transformants were screened by restriction digest and confirmed by DNA sequencing.

2.4.3 Immunofluorescence of HLA-A2-GFP and dicysteine containing variants

HeLa cells were transfected using Fugene HD (Roche) according to the recommended protocol. The following day, #1.5 round glass coverslips were coated with alcian blue by submerging the coverslips in a solution of 1% (w/v) alcian blue in water. The solution (and coverslips) was boiled under close supervision using microwave heating for 20 seconds, after which the solution was left to cool at room temperature for 10 min. After cooling, the alcian blue solution was recovered, and the coverslips washed extensively in deionized water. The coverslips were sterilized by submersion in 75% ethanol

Primer Name	Note	Sequence (5′ - 3′)
HLAA2R5	Reverse Seq	CAGCGTGGTGAGTCATATGCGTTT
HLAA22F	Forward Seq	AGAGGATGGAGCCGCGGG
HLAA23F	Forward Seq	CGCGGGGAGCCCCGCTTCAT
R89_A93	QuickChange	GAGTATTGGGACGGGGAGACATGCAAAGTGAAGTGCCACTCACAGA
R89_A93_R	"	TCTGTGAGTGGCACTTCACTTTGCATGTCTCCCCGTCCCAATACTC
K90_A93	"	GGGACGGGGAGACACGGTGCGTGAAGTGCCACTCACAGACTCACC
K90_A93_R	"	GGTGAGTCTGTGAGTGGCACTTCACGCACCGTGTCTCCCCGTCCC
Q179_A182	"	GGCCCATGTGGCGGAGTGCTGGAGATGCTACCTGGAGGGCACG
Q179_A182_R	"	CGTGCCCTCCAGGTAGCATCTCCAGCACTCCGCCACATGGGCC
Q179_Y183	"	CATGTGGCGGAGTGCTGGAGAGCCTGCCTGGAGGGCAC
Q179_Y183_R	"	GTGCCCTCCAGGCAGGCTCTCCAGCACTCCGCCACATG
g72c	Add NgoMIV site	CCAGACCTGGGCCGGCTCTCACTCC
g72c_R	"	GGAGTGAGAGCCGGCCCAGGTCTGG
NC4-Tag		CAGCGCCGGCTTTTTAAATTGCTGCCCCGGGTGCTGCATGGAACCGGGAAGTAGTGGAAGTAGTGCCGGCGAGC
NC4-Tag_R	"	GCTCGCCGGCACTACTTCCACTACTTCCCGGTTCCATGCAGCACCCGGGGCAGCAATTTAAAAAGCCGGCGCTG

Table 2.2: Oligonucleotides used in this study. All are listed 5′-3′. The 'R' indicates the oligonucleotide is the reverse-complement of another entry with the same designation, or designed to sequence in the 3′-5′ direction.

in water, and allowed to dry on a kimwipe in a cell culture hood. After drying the coverslips were placed in a 12 well tissue culture plate. Transfected HeLa cells were lifted from their culture flask by treatment with trypsin, diluted in media, and approximately 50,000 transfected HeLa cells were added to each coverslip-containing well and allowed to adhere overnight. The following day, cells were stained by incubation in FACS buffer (Hanks Balanced Salt Solution, HBSS, Gibco Cat# 14025, with 0.5% (w/v) bovine serum albumin added) containing a 1:500 dilution of BB72-PE for 1 hr. After labeling, cells were washed extensively in labeling buffer, and then lightly fixed and nuclei labeled by submersion in fixation buffer (4% paraformaldehyde in PBS) containing 300 nM 4',6-diamidino-2-phenylindole (DAPI) for 5 min at RT, after which they were washed and mounted on glass slides using GelMount (Biomeda) or FluoroGel (Electron Microscopy Sciences). The slides were imaged using a 40X LD-NeoFluar objective on a Zeiss Axiovert 200M. Images for GFP, PE and DAPI fluorescence were aquired using Zeiss filter sets 44, 43, and 49, respectively, on a MRm camera and analyzed using ImageJ.

2.4.4 Flow cytometric analysis of surface-expression of HLA-A2-GFP and di-cysteine containing variants

HeLa or Buf1280 growing in plastic tissue culture treated plates were transfected using Lipofectamine2000 (Invitrogen) according to the manufacturers instructions. After growing overnight, the cells were washed twice with PBS (without calcium or magnesium) before lifting from the growth substrate by incubating at 37°C for 30 min in 2 mL PBS supplemented with 1 mM each EDTA and EGTA. The cells were collected by centrifugation (500 x *g*, 7 min) and resuspended in FACS buffer. The cells were then placed in a round bottom 96 well microtitre plate and labeled by incubation for 1 hr at 4°C in 150 μL of FACS buffer containing a 1:500 dilution of BB7.2-PE. After labeling, the cells were washed twice in FACS buffer, and resuspended in 500 μL FACS buffer containing 1 μg/mL propidium iodide. The cells were maintained on ice until analysis by flow

cytometry (FACSCalibur, BD). Histograms were constructed in FlowJo (TreeStar) after excluding dead cells by gating on forward and side scatter, and selecting a propidium iodide negative population. Channel compensations were conducted at the time of acquisition using control samples containing only a single fluorophore. In some cases, the mean fluorescence intensity (MFI) is shown. This value is the geometric mean of at least 10,000 cells falling within live gates defined by forward scatter, side scatter, and PI fluorescence. In some cases, these values are gated on GFP+ cells, in which case the number of cells analyzed before GFP gating was $\geq$10,000.

2.4.5 ReAsH staining of tetracysteine-tagged HLA-A2-GFP

HeLa cells grown on glass bottom dishes (MatTek) were transfected with C4-HLA-A2-GFP using Fugene according to the manufacturer's instructions. After overnight expression, the cells were washed twice with HBSS and incubated for 20 min at 37°C in 200 μL ReAsH labeling solution containing 1 μM ReAsH, 10 μM EDT, 10 mM tributylphosphine, 1 mM $CaCl_2$ and 1 mM $MgCl_2$ in PBS (Gibco Cat# 14190). The cells were then washed twice with 2mL PBS and incubated for 20 min at 37°C in 2 mL ReAsH wash solution containing 100 μM EDT, 10 mM tributylphosphine, 1 mM $CaCl_2$ and 1 mM $MgCl_2$ in PBS. The cells were washed twice with PBS and incubated for 10 min at room temperature in 2mL HBSS containing 300 nM Hoescht 33342 before imaging. Images were acquired using a 40X LD-NeoFluar objective on a Zeiss Axiovert 200M. Images for GFP, ReAsH and Hoescht fluorescence were aquired using Zeiss filter sets 44, 43, and 49, respectively, on a MRm camera and analyzed using ImageJ. To better visualize ReAsH labeling of endosomes, some images were enhanced using the 'unsharp mask' filter, in which the image is blurred and the blurred image subtracted from the parent image. This has the effect of enhancing bright spots and decreasing background.

2.4.6 Surface stabilization of dicysteine containing HLA-A2 variants by serum

Buf1280 or HeLa cells transfected with HLA-A2-GFP or cysteine containing variants were grown in RPMI-1640 + 10% FBS overnight. The following day, the cells were washed twice with 2 mL HBSS before incubating (at 37°C) in either HBSS, HBSS + 10% FBS, RPMI-1640 (serum free media) or RPMI-1640 + 10% FBS (media + serum), for the indicated length of time. The cells were them washed twice with 2 mL PBS and detached from the plate using 2 mL PBS supplemented with 1mM each EDTA and EGTA. The cells were diluted into FACS buffer, collected by centrifugation, and stained for surface HLA-A2 as described in §2.4.4.

To determine the rate of disappearance of surface HLA-A2, Buf 1280 cells grown in 6 well tissue culture plates were transfected with HLA-A2-GFP and allowed to express the protein overnight in RPMI-1640 + 10% FBS. At selected times the following day, individual wells were washed twice with 2mL HBSS, and the incubated further for varying amounts of time in HBSS without serum. Following incubation the cells were detached from the plate using 2 mL PBS supplemented with 1mM each EDTA and EGTA. The cells were diluted into FACS buffer, collected by centrifugation (500 x g, 7 min) and stained for surface HLA-A2 as described above.

2.4.7 Surface stabilization of dicysteine containing HLA-A2 variants by peptide

Buf1280 cells transfected with HLA-A2-GFP or cysteine containing variants were grown in RPMI-1640 + 10% FBS overnight. The following day, the cells were washed twice with 2 mL HBSS before incubating for at least 8 hr in either HBSS or HBSS supplemented with 10 μM YLLCCNYKL. The cells were them washed twice with 2 mL PBS and detached from the plate using 2 mL PBS supplemented with 1mM each EDTA and EGTA. The cells were diluted into FACS buffer, collected by centrifugation, and stained for surface HLA-A2 as described in §2.4.4.

Chapter 3

Discrete Arginine Topologies Direct Cytosolic Access

3.1 Background

Protein drugs–also known as biologicals–represent a rapidly expanding class of therapeutic molecules[105]. However, their actions are limited primarily to extracellular targets[106, 107] because the size and composition of most polypeptides and proteins do not facilitate their uptake into mammalian cells[108]. It has been known for over 40 years that addition of cationic charges to a peptide or protein can aid transport into cells[109, 110, 111], and many reports have demonstrated the utility of appending basic sequences derived from the HIV Trans-Activator of Transcription (**Tat**), *D. melanogaster* Antennapedia, or simply polyarginine (for example, **Arg8**) to peptides to increase their cytoplasmic access[112, 113, 114, 115, 116, 117]. Certain highly positively charged proteins also possess cell penetrating properties[118, 119], but it remains unknown how one can leverage these examples to control the precise entry pathway or enhance the determinants for uptake.

Three contrasting mechanisms for cytosolic entry of cationic proteins and polypeptides have been proposed. The first (ion pair-guided passive diffusion) posits that guani-

dinium side chains on the polypeptide form hydrogen bonds with cell surface phospholipids creating neutral ion pairs that passively diffuse across the plasma membrane[120, 121]. The second model (endosomal release), asserts that endocytosis is a major portal through which cationic proteins and polypeptides enter the cell[106, 122, 123], followed by endosomal release. Previous investigations have attempted to distinguish between these two models by blocking endocytosis, via thermal[115], pharmacologic[116, 123], or genetic means[116, 124]. The interpretation of these experiments is complicated, however, by differences in protein/polypeptide concentration and analytical method. Recently, a third mechanism for cytosolic entry was described: at concentrations $\geq$ 10 μM incubation of living cells with cationic proteins/polypeptides leads to the formation of nucleation zones[125] that transiently disrupt membranes[126]. This results in the spontaneous release of peptide into the cytosol. In the presence of drugs that inhibit endocytosis, incubation of cells at lower concentrations ($\leq$ 5 μM) of peptide prevents cytoplasmic access[116], implying that, at low concentrations, the molecules studied cannot diffuse through the plasma membrane. The many studies using microscopy to examine cells fixed by treatment with formaldehyde or methanol[127] must be reevaluated in light of evidence that the fixation process releases fluorescently labeled peptides from endosomes, an artifact not observed during microscopic examination of living cells[127]. Finally, the high intensity light used for illumination during microscopy can itself facilitate the redistribution of fluorescently labeled peptides from endosomes to cytoplasm[128]. Thus, whether, when, and how these cationic molecules escape endocytic vesicles to access the cytosol remain unanswered questions.

Attempts to identify and improve upon the structural determinants of cell permeability are complicated by the above experimental details as well as the fact that neither **Tat** nor **Arg8** possesses a defined fold. Miniature proteins are a family of small (36-aa), well-folded polypeptides that adopt a characteristic hairpin fold consisting of axially packed α- and PPII helices[129, 130]. Miniature proteins identified through both rational design[15, 17] and molecular evolution[131, 132, 19] can modulate protein function by

inhibiting protein interactions[132, 18]; both loss of function and gain of function activities have been observed[132, 19, 18]. We reported previously that minimally cationic miniature proteins containing between 2 and 6 arginine residues embedded within the α- or PPII helix were taken up by mammalian cells in culture more efficiently than **Tat** or **Arg8**[133, 134]. Now, we turn to investigate whether miniature proteins containing arginine reach the cytoplasm.

3.2 Results

To learn more about the structural determinants of cytoplasmic access, we designed a set of miniature proteins that differed in the number and density of α-helical arginine side chains, and tracked their passage into the cell. Using low concentrations (1 μM) of variants conjugated to a fluorophore, we found that a minimum of 4 α-helical arginines was required for uptake, and cell uptake was enhanced when the arginines were clustered on the same α-helix face. Next, a novel and rapid assay for evaluating cytoplasmic access revealed that of four cationic miniature proteins taken up by cells, only one reaches the cytosol. This miniature protein, which we termed **5.3**, possesses a distinct array of five dispersed α-helical arginines on three helical faces. Direct comparisons show that after treatment for as little as 30 min, the cytoplasmic concentration of **5.3** reaches levels as high or higher than that achieved by **Tat** or **Arg8**. Experiments monitoring the degradation of proteins that reach the cytosol show that 5.3 displays remarkable stability compared to **Tat** or **Arg8**, cationic peptides that do not posses a stable fold. These experiments demonstrate that discrete arginine arrangements embedded within a well-folded miniature protein can direct cytosolic access.

3.2.1 Miniature Protein Design

To examine the effect of charge density and orientation on cell uptake, I prepared a variety of miniature proteins, using avian pancreatic peptide (aPP, structure shown in Fig.

3.1a) as a scaffold[130, 129], that contained variety in the number and positioning of positively charged arginine substitutions. The stability of the aPP fold[135] allows variation to be introduced into solvent exposed residues without compromising structure. The eight miniature proteins, as well as the wild-type sequence, together contained between one and six arginine residues at various positions on the solvent-exposed α-helical surface of the hairpin fold (Fig. 3.1b). These molecules also contained two arginines near the C-terminus (Fig. 3.1c).

To ensure the arginine residues were displayed as intended, circular dichroism (CD) spectra of the miniature protein variants were obtained (Fig. 3.2). Seven of these cationic miniature proteins were characterized by CD spectra at 37°C that were virtually indistinguishable from that of the parent molecule lacking additional arginines, aPP (Fig. 3.2a). The CD spectra of six were temperature-dependent with cooperative transitions between 49 and 67°C (Fig. 3.2b,c) suggesting that they each retained a stable and characteristic hairpin fold[130]. Miniature protein **6.3**, containing the greatest number of arginine substitutions (6), showed reduced ellipticity at 222 nm and 208 nm (Fig. 3.2a), along with a reduced T_m of 33°C (Fig. 3.2b,c); **6.3** was not studied further. For the remaining molecules, these data suggest that arginine substitution does not significantly alter miniature protein secondary structure, and that Fig. 3.1b accurately represents the arginine side chain arrangement in miniature proteins **2.1**, **2.2**, **3.2**, **4.2**, **4.3**, **5.2**, and **5.3**.

3.2.2 Cell Uptake

Initially we used flow cytometry to assess the influence of arginine number and orientation on miniature protein uptake. In this technique cells are passed through a capillary into a detection chamber illuminated by one or several lasers. Appropriately placed photodetectors fitted with bandpass filters detect fluorescent signals. This allows the amount of fluorescence associated with individual cells to be accurately and rapidly determined. In preliminary studies, we evaluated molecules labeled with fluorescein on

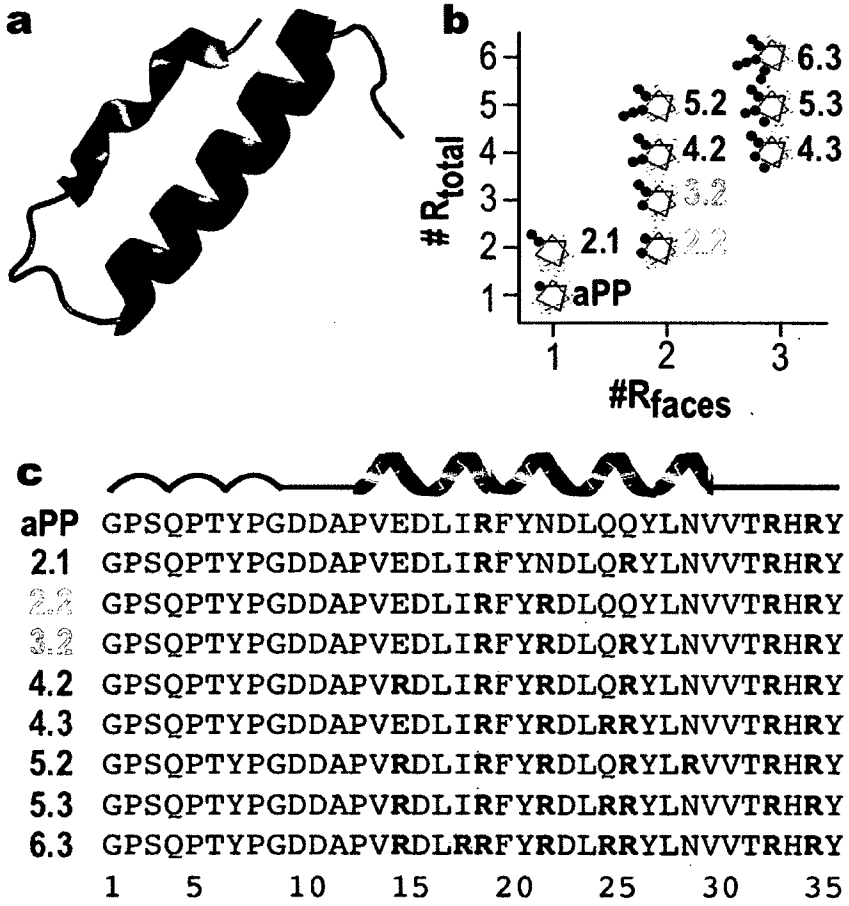
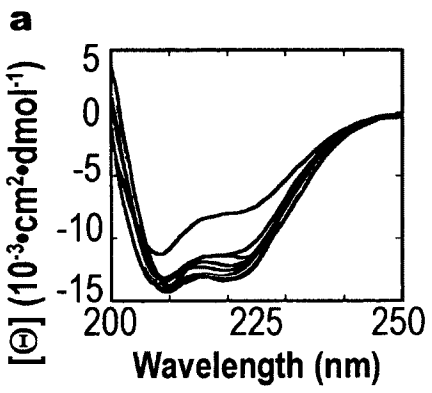


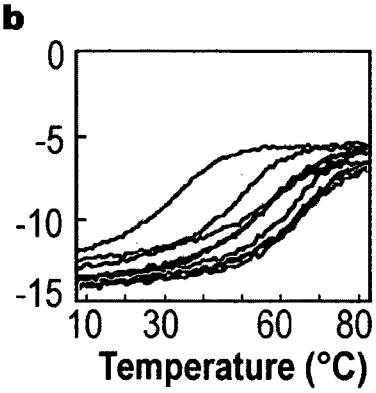
Figure 3.1: Description of miniature protein structures and variation in arginine display. (**a**) The structure of avian pancreatic polypeptide (PDB:1PPT) contains both α- and polyproline-type helices, which are arranged in a hairpin. The interface is formed *via* hydrophobic packing. The focus of this study was on determinants that could be placed on the α-helix. (**b**) A plot of the relationship between $\#R_{total}$ (the number of α–helical arginine residues) and $\#R_{faces}$ (the number of α–helical faces on which these arginines are displayed) for each aPP variant. The location of each α-helical arginine residue is represented by a blue circle on the helical wheel. (**c**) Amino acid sequences of peptides shown in **b**. *[This data assembled with Betsy Smith.]*

their C-termini (as denoted with superscript F). In our initial studies, we included a trypsin wash just before analysis via flow cytometry. Treatment of HeLa cells with 5 μM **aPP**^{F}, **2.1**^{F}, **2.2**^{F}, or **3.2**^{F} resulted in only small increases (< 3 fold) in cell fluorescence, while treatment with **4.2**^{F}, **4.3**^{F}, **5.2**^{F}, or **5.3**^{F} resulted in increases in cell fluorescence between 7 and 40 fold (Fig. 3.3a). We therefore chose to focus on miniature proteins showing significant uptake.

To ensure that increased uptake was a property of peptide sequence and not of fluorophore location or identity, we confirmed these results by synthesizing analogs labeled instead on their N-termini with tetraethyl rhodamine sulfate (denoted with superscript R), a dye with several desirable properties including resistance to photobleaching and an emission spectrum unaffected by pH changes and far from the autofluorescence spectrum of cells[57]. As found for miniature proteins labeled with fluorescein, rhodamine labeled miniature proteins containing four or five α-helical arginines were taken up efficiently (Fig. 3.3b). Comparison of the relative fluorescence signal of cells treated with peptides sharing the same sequence but labeled with different fluorophores reveals a striking similarity, suggesting uptake is largely influence by the nature of the peptide, and not purely a characteristic of the fluorophore tag. We found that incubating HeLa cells for longer time periods increased the fluorescence signal associated with cells (Fig. 3.3c), suggesting that peptides were being internalized in an ongoing, rather than stochastic, process. Comparison of the fluorescent signal after 90 min vs 30 min reveals that the rate of uptake is highest for miniature proteins containing clustered arginines (**4.2**^{R} and **5.2**^{R}) whose fluorescence intensities increase by 160% and 281% during this time period, and is lower for **Tat**^{R} and **Arg8**^{R}.

Importantly, flow cytometry places no restrictions on *where* a fluorescent signal originates. Fluorescent molecules on the cell's surface, within the cytoplasm, or within intracellular vesicles all contribute to the total cellular fluorescence. To distinguish molecules within cells, from those that remain on the cell's surface, previous researchers have attempted to decrease surface binding through lowering pH[136, 137], adding competing





c

Molecule	T_m (°C)
aPP	64
2.1	65
2.2	65
3.2	67
4.2	56
4.3	60
5.2	55
5.3	49
6.3	33

Figure 3.2: Circular dichroism (CD) of miniature proteins used in this study. (**a**) Wavelength dependent CD spectra of peptides (25 μM) in PBS at 37°C. (**b**) Temperature dependent ellipticity at 222 nm of peptides shown in a. (**c**) Calculated T_m of thermal melts shown in **b**. *[This data collected by Betsy Smith.]*

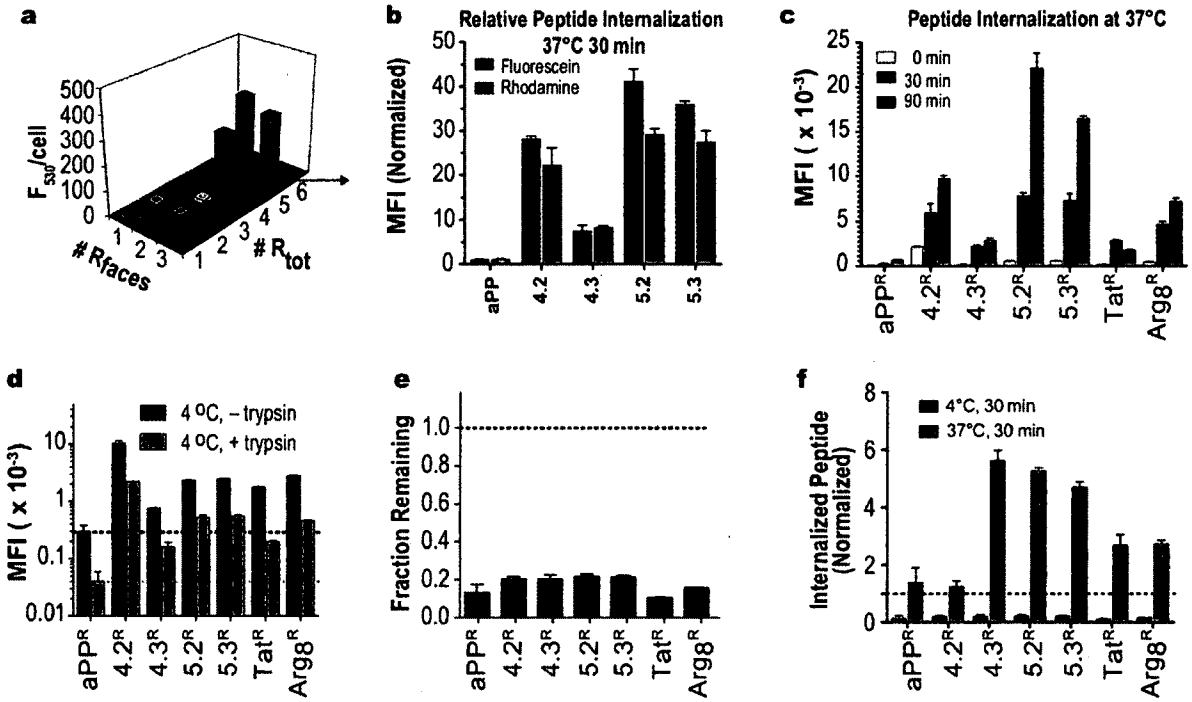


Figure 3.3: Cell uptake of cationic miniature proteins. (**a**) A plot of mean cellular fluorescence at 530 nm of HeLa cells treated with fluorescently tagged miniature proteins (5 μM, 30 min). (**b**) Comparison of the relative uptake of fluorescein and rhodamine labeled miniature proteins sharing the same sequence. HeLa cell treated as in **a**, or with 1 μM rhodamine labeled miniature proteins for 30 min at 37°C. The mean fluorescence intensity (MFI) is normalized to the value for aPP. (**c**) Time dependent internalization of cationic miniature proteins. HeLa cells treated with 1 μM rhodamine labeled miniature proteins at 37°C for 30, or 90 min or treated at 4°C ('0 min') before analysis. (**d**) Cellular fluorescence after treatment with 1 μM rhodamine labeled miniature proteins in the absence of endocytosis and removal by trypsin treatment. HeLa cells were treated for 30 min at 4°C were then treated with trypsin (0.05%, 10 min, 37°C) or PBS before washing and analysis by flow cytometry. (**e**) Fraction of cell-associated fluorescence (**d**, black bars) remaining after trypsin treatment (**d**, red bars). (**f**) Internalization efficiency of rhodamine labeled miniature proteins. *[The data in part* **a** *was collected by Betsy Smith]*

binding substrate such as heparan sulfate[138], or by degrading surface-bound protein with trypsin[138, 127]. I therefore verified that trypsin treatment removed miniature proteins that were bound to the cell surface. Membrane traffic was arrested by incubating cells at 4°C[139, 140] for 15 min prior to and during a 30 min treatment with 1 μM $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, $\mathbf{Tat}^R$, $\mathbf{Arg8}^R$ or $\mathbf{aPP}^R$ as a negative control. After incubation, cells were washed three times with PBS and incubated with trypsin (0.05% at 37°C for 15 min) or PBS (as a control) before analysis by flow cytometry (Fig. 3.3d). Cationic miniature proteins $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$ and $\mathbf{5.3}^R$ bound to cells between 2.7-fold and 35-fold more than $\mathbf{aPP}^R$ and in some cases ($\mathbf{5.2}^R$ and $\mathbf{5.3}^R$) to an extent comparable to $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$. For all peptides, trypsin treatment decreased the fluorescent signal between 77-89% (Fig. 3.3e). Because peptides and proteins that enter the cell are inaccessible to trypsin added to the culture media[113], these data confirm that at 4°C, incubation of cells with 1 μM cationic miniature protein leads to little if any cell uptake, and confirms that the vast majority of material bound to the cell surface is degraded and/or effectively removed by trypsin treatment.

Our finding that incubating cells in the cold substantially decreases the amount of internalized fluorescence signal indicates that at 4°C neither cationic miniature proteins nor **Tat** or **Arg8** can enter the cell by diffusing through the plasma membrane. Two hypotheses could explain this result. First, at 4°C the lipid character of the membrane may be altered is such a way as to become less permeable to the peptides studied. Alternatively, the internalization of membrane is required for cell uptake, a process arrested at 4°C. These theories are not mutually exclusive. Nevertheless, this finding also provided an opportunity to compare the efficiency of internalization, defined herein as the amount of peptide internalized relative to that bound to the cell surface. In the absence of endocytosis (4°C), the amount of surface bound peptide was quantified using flow cytometric analysis of cells that had been treated with 1 μM $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, $\mathbf{Tat}^R$, $\mathbf{Arg8}^R$ or $\mathbf{aPP}^R$, but not with trypsin(Fig. 3.3d). Values obtained after incubating cells at either 4°C or 37°C (Fig. 3.3c) were normalized to this value. As can be seen in figure 3.3f, $\mathbf{aPP}^R$

and 4.2^R are internalized inefficiently, showing fluorescence values after 30 min roughly equal to the amount originally bound to their surface. By contrast, 4.3^R, 5.2^R and 5.3^R are each internalized efficiently, incorporating more than 4 times the amount of peptide originally bound to their surface. **Tat**^{R} and **Arg8**^{R} show intermediate values, revealing that the display of arginine side chains on cationic miniature proteins can be engineered to induce highly efficient uptake by cultured cells.

These data show that cell uptake depends on arginine density and number; miniature proteins containing four arginine residues clustered on two helical faces were taken up to a greater extent than those containing four arginines on three helical faces (Fig. 3.3a,b). Molecules containing five α-helical arginines were taken up to a similar extent irrespective of density, revealing that among these molecules the impact of arginine arrangement was smaller. Consistent with our previous work[141], the cationic miniature proteins 5.2^R and 5.3^R are taken up to a greater extent and with higher efficiency than **Tat**^{R} and **Arg8**^{R}, despite the fact that they possess twice the mass and fewer (7 rather than 8) positive charges.

3.2.3 Cationic miniature proteins do not cause plasma membrane disruption.

Three limiting models have been invoked to explain the trafficking of cationic peptides and proteins across the plasma membrane. The first posits that high local concentrations of positively charged peptides cause membrane disruption[125]. This is followed by sudden release of peptide into the cell and triggers a membrane repair response[126]. A second model invokes ion-pair guided passive membrane diffusion[121, 120]. The final model invokes endocytosis followed by endosomal release (*vide infra*). Membrane disruption can alter cell physiology leading the leakage of cytoplasmic contents and cell death[142]. Therefore, we first confirmed that the integrity of the cell membrane is not disrupted in the presence of 1 μM cationic miniature protein.

To confirm that cationic miniature proteins do not cause membrane disruption, we

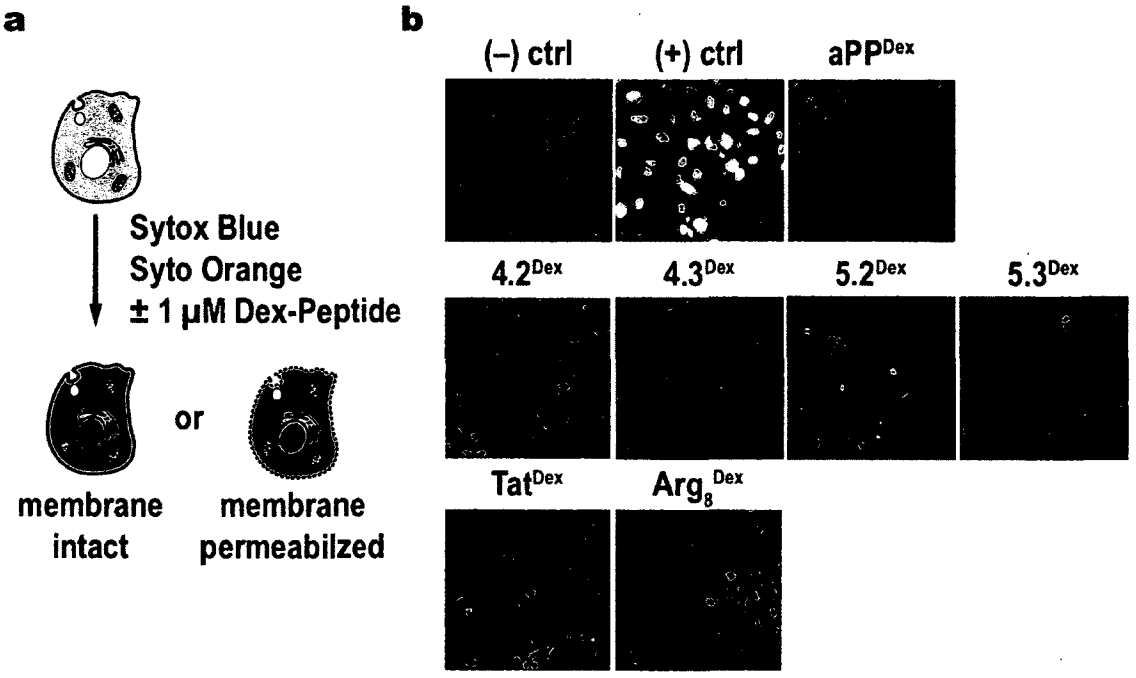


Figure 3.4: Absence of membrane permeabilization after treatment with Dex-labeled cationic miniature proteins. HeLa cells were incubated with 1 µM of the indicated dexamethasone labeled miniature protein or peptide in the presence of 200 nM Syto Orange (a cell permeable dye) and 200 nM Sytox Blue, a dye that is not cell permeable but binds to nuclear DNA when cell membranes are disrupted. Saponin (10 µg/mL) a detergent that permeabilizes mammalian cell membranes was used as a positive control. (**a**) Schema showing experimental design; cells with intact membranes exclude Sytox blue and are red only, cells with permeabilized membranes are red and blue. (**b**) Images of cells treated as described in **a** with the indicated peptides. Very few show blue signal in the nucleus, indicating intact cell membranes.

treated HeLa cells with peptides in the presence of two dyes, one that is cell permeable (Syto Orange), and one that is not (Sytox Blue). In the absence of membrane disruption, cell bodies are labeled with the permeable dye and exhibit red fluorescence. No blue fluorescence is observed from cells with intact membranes, however membrane disruption allows for a high affinity interaction between Sytox Blue and DNA leading to bright fluorescence within the nucleus (see Fig. 3.4a). As can be seen in figure 3.4b, HeLa cells treated in media, indicated as (–) CTRL, show only red fluorescence, but after treatment with a saponin, a detergent capable of permeabilizing mammalian cell membranes, bright blue fluorescence becomes visible within the nucleus. To test whether treatment with 1μM cationic miniature proteins would lead to similar membrane compromise, we treated HeLa cells with miniature proteins or peptides as indicated under analogous conditions in the presence of Sytox Blue and Syto Orange for 30 min at 37°C. As can be seen in figure 3.4b, very few cells show blue fluorescent nuclei, indicating that dexamethasone labeled cationic miniature proteins do not lead to membrane disruption. This result suggests that at low concentrations and under these condition, neither cationic miniature proteins nor **Tat** or **Arg8** enter cells through this mechanism.

3.2.4 Cationic Miniature Proteins Reach The Cytoplasm

In many cases, cationic proteins/polypeptides showing cellular uptake by flow cytometry remain trapped in endosomes and fail to reach the cytosol[128]. Previous work has shown that the cytosolic localization of a peptide- or peptoid-dexamethasone (Dex) conjugate can be detected using a glucocortocoid receptor-driven luciferase assay[143]. Interaction of a Dex-labeled peptide or peptoid with the ligand-binding domain of the glucocorticoid receptor (GR) in the cytosol induces transcription of luciferase that is detectable in cell lysates 40-48 hours later. We sought to evaluate cytosolic localization on a more rapid time scale, and took inspiration from reports that dexamethasone treatment of cells in culture led to rapid (15 min) translocation of a GR-green fluorescent protein

fusion (GR-GFP) from the cytosol to the nucleus[144]. If this rapid translocation was also observed in the presence of peptide-dexamethasone conjugates, then live cell microscopy could provide a rapid assay for cytosolic localization in living cells.

To validate this approach, HeLa cells were transfected for 12 - 18 h with the gene encoding GR-GFP, and then incubated for 30 min with between 0 and 10 μM dexamethasone (Fig. 3.5a). When dexamethasone was absent from the incubation media, these cells exhibited GFP signal throughout the cytoplasm and the nucleus (Fig. 3.5e). Addition of dexamethasone to a final concentrations between 3 nM and 10 μM led to a concentration-dependent decrease in the cytosolic GFP signal and a concomitant increase in the nuclear GFP signal (Fig. 3.5a). We quantified these changes using the automated image processing package CellProfiler[145] (Fig. 3.5b,c) to measure the GFP signal in the nucleus as well a 2 μm (40 pixel) region surrounding region of the cytosol. I report the ratio of the median GFP intensity in the nucleus and the surrounding region as a 'translocation ratio'[1] (TR, Fig. 3.5c). TR values near 1 indicate equivalent intensity between the nucleus and the surrounding region. Treatment of HeLa cells with 1 μM dexamethasone for 30 min resulted in an increase in the translocation ratio from 1.07 ± 0.02 to 3.93 ± 0.14 ($p = 1.92 \times 10^{-22}$, vs untreated cells, two-tailed t-test), roughly 90.5% of value achieved with a 10-fold higher concentration(Fig. 3.5d). We therefore chose this concentration for subsequent studies.

When HeLa cells transfected with GR-GFP were treated for 30 min with 1 μM $\mathbf{aPP}^{Dex}$, virtually no change in GR-GFP localization (Fig. 3.5e) or the calculated translocation ratio (1.41 ± 0.09, vs 1.25 ± 0.04, $p = 0.1473$, two-tailed t-test, see also Fig 3.6b) was observed. Among HeLa cells that were treated with media alone, the translocation ratio was not significantly correlated with the level of expression of GR-GFP ($R = 0.21$, Fig. 3.5f). These data are consistent with the observation that $\mathbf{aPP}^{R}$ fails to enter cells and indicate that simply adding the dexamethasone label is not sufficient to confer upon an

[1]Others (e.g., Ding et al. (1998) J Biol Chem 273, 28897–28905) have reported similar values, which they have named nuclear/cytoplasmic ratio, however because this term could be misconstrued to indicate the relative concentration of GFP within these compartments, I have used the term translocation ratio instead.

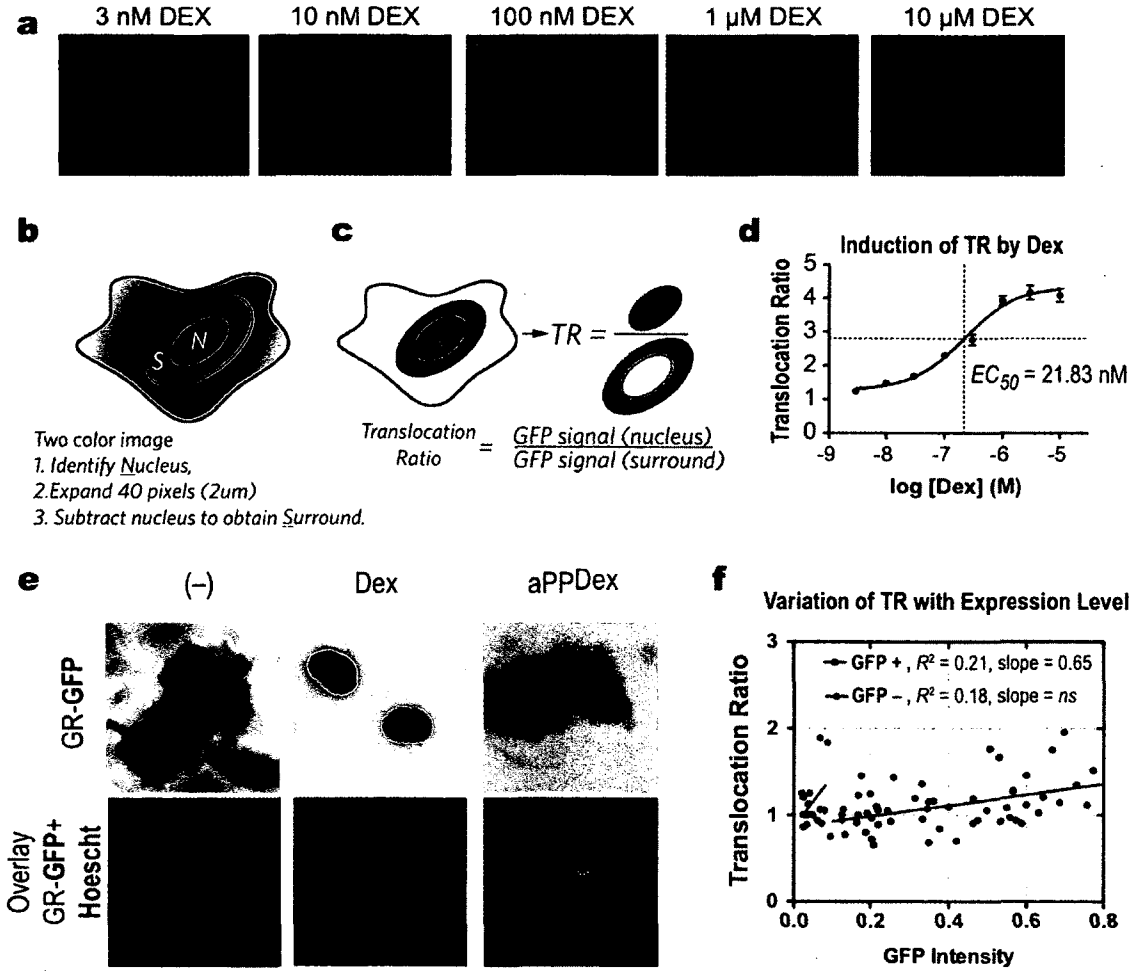


Figure 3.5: Dose dependent increase in translocation ratio following exposure to dexamethasone. (**a**) Schematic of quantitation method. Cell images were examined using CellProfiler to create a region containing the nucleus and a surrounding zone. The ratio of the median intensity in these two regions is reported as the translocation ratio. (**b**) Representative images of HeLa cells transfected with GR-GFP and treated with various concentrations of dexamethasone as indicated. Red outlines show regions defined as nucleus and surround. (c) Translocation ratio observed plotted against the concentration of dexamethasone added. Fitting these points to a hyperbolic curve yields an EC_{50} = 21.83 nM and a Hill coefficient of 0.97.

otherwise cell impermeable peptide the ability to reach the cytoplasm[143]. These data also support the use of the translocation ratio as a useful measure of the relative ability of Dex-labeled compounds to reach the cytoplasm of target cells.

I next evaluated the extent to which Dex-labeled miniature proteins, as well as **Tat** and **Arg8**, induce the nuclear translocation of GR-GFP (Fig. 3.6). HeLa cells transfected with GR-GFP were treated for 30 min with 1 μM $\mathbf{4.2}^{Dex}$, $\mathbf{4.3}^{Dex}$, $\mathbf{5.2}^{Dex}$, $\mathbf{5.3}^{Dex}$, $\mathbf{Tat}^{Dex}$, or $\mathbf{Arg8}^{Dex}$ and the translocation ratio calculated. Treatment of HeLa cells with 5.3Dex led to a large increase in translocation ratio (3.1 ± 0.1) both compared to an untreated sample (–) ($p = 2.68 \times 10^{-48}$, ANOVA with Bonferroni post-test) or one treated with $\mathbf{aPP}^{Dex}$ ($p = 1.96 \times 10^{-16}$). Treatment of HeLa cells with $\mathbf{4.2}^{Dex}$, $\mathbf{4.3}^{Dex}$, or $\mathbf{5.2}^{Dex}$ led to a small or absent increase in nuclear GFP signal, and translocation ratios (1.18 ± 0.04, 1.92 ± 0.10, 1.77 ± 0.11 respectively) that were not significantly different from cells treated with $\mathbf{aPP}^{Dex}$. We also measured the nuclear accumulation of GR-GFP after treatment with $\mathbf{Arg8}^{Dex}$ or $\mathbf{Tat}^{Dex}$, which in both cases was increased over control samples treated with $\mathbf{aPP}^{Dex}$ ($\mathbf{Arg8}^{Dex} = 2.19 \pm 0.13$, $p = 0.0053$; $\mathbf{Tat}^{Dex} = 2.85 \pm 0.10$, $p = 2.54 \times 10^{-9}$), and comparable to cells treated with $\mathbf{5.3}^{Dex}$. These comparisons are illustrated in figure 3.6b.

The above data could be explained in several trivial ways. One explanation is that the concentrations of Dex-labeled cationic miniature proteins and peptides used here could lead to permeabilization of the cell membrane and direct access of the peptides studied to the cytoplasm. However I previously verified that the plasma membrane retains the ability to exclude the small organic dye Sytox Blue (Fig 3.4), making this possibility unlikely. Second, the peptide could be rapidly degraded, releasing free dexamethasone into the cell and surrounding media, activating the GR and leading to its nuclear accumulation. Third, and perhaps less likely, by covalently linking dexamethasone to the cationic miniature protein, a high affinity molecule is created, capable of activating the GR at extremely low concentrations. I tested these final two possibilities directly by assaying the proteolytic stability of the miniature protein backbone and by measuring the ability

of dex-labeled cationic miniature proteins to compete with dexamethasone for the ligand binding site of recombinantly expressed human glucocorticoid receptor *in vitro*.

To verify that the affinities of dexamethasone labeled cationic miniature proteins and peptides for the human glucocorticoid receptor in vitro were similar to each other, we[2] used a commercially available validated glucocorticoid screening assay, PolarScreen Green®. In this assay, recombinant human glucocorticoid receptor is mixed with a fluorescently labeled steroid analog (Fluormone), that possesses a high affinity for the mineralocorticoid binding site. In the absence of competitor molecules, Fluormone binds to GR and exhibits high polarization. As increasing concentration of competetor molecule is added, Fluormone is displaced from the binding site. As its effective molecular size is now reduced, the molecule tumbles in solution more quickly, and exhibits a polarization value that is decreased. Fitting these data to a hyperbolic curve reveals the IC_{50}, the concentration of competetor molecule that results in a half-maximal decrease in the polarization value of Fluormone.

Addition of cationic residues within the α-helix of the aPP scaffold resulted in a trend of decreasing ability of Dex-labeled cationic miniature proteins to compete for the GR binding site (Fig 3.7). The affinity of $\mathbf{aPP}^{Dex}$ for the GR (IC_{50} = 1.0 nM, 95% CI = [0.6, 1.7]) was similar to dexamethasone itself, (IC_{50} = 0.85 nM, 95% CI = [0.5, 1.4]), whereas the affinity of $\mathbf{5.3}^{Dex}$ was lower (IC_{50} = 35.9 nM, 95% CI = [25.1, 51.3]). Miniature proteins $\mathbf{4.2}^{Dex}$, $\mathbf{4.3}^{Dex}$, and peptides $\mathbf{Tat}^{Dex}$ and $\mathbf{Arg8}^{Dex}$ gave intermediate values (between 2.7 – 8.6 nM), all slightly poorer than unconjugated dexamethasone (see Fig 3.7c-h). Because the apparent affinity of $\mathbf{5.3}^{Dex}$ is *lower* than the other molecules studied in this panel, its increased ability to activate the glucocorticoid receptor in cells must invoke either a non-competetive mechanism or a higher effective concentration within the cytoplasm, or both.

[2]This work was completed with the assistance of Justin Holub

3.2.5 Cell Penetrating Cationic Miniature Proteins Reach The Cytosol Intact

A trivial explanation for the results obtained above is that cationic miniature protein **5.3** as well as **Tat**, and **Arg8** are degraded by endosomal proteases, releasing the dexamethasone label, which, by itself, can efficiently cross cell membranes and induce the nuclear localization of GR-GFP. Thus, I sought to confirm that the peptide backbone of miniature protein **5.3** is not degraded under the conditions in this assay, that is to say, within the 30 min incubation period after the peptides are added to HeLa cells and before they are analyzed by epifluorescence microscopy. To examine possible metabolism of **5.3**, HeLa cells were treated with 1 μM rhodamine labeled cationic miniature protein under conditions analogous to those used for flow cytometry above (Fig. 3.3). The cells were then pelleted, washed with PBS, and lysed using a detergent buffer. The insoluble material was cleared from the lysate by a brief centrifugation (10 min, 10000*g*) after which the supernatant containing soluble cell lysate was diluted into 15% ethanol in water and injected onto an HPLC fitted with a C8 (100 x 2.1 mm, 100Å pore) collum. Because we expected the concentration of this solution to be low, we employed a high sensitivity Shimadzhu HPLC fitted with a fluorescence detector capable of sensing the presence of attomoles of rhodamine. The fluorescence HPLC traces obtained are shown in figure 3.8. As can be seen in panels **a** and **c**, the fluorescence chromatogram of miniature proteins obtained from the cell lysate are virtually unchanged from the chromatogram on an input sample. By contrast, the chromatograms of $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$ show an increased number of peaks in chomatogram obtained from lysate as compared to the chromatogram obtained from the input sample.

It is likely in the above experimental set up that some of the peptide obtained after cell lysis was bound to extracellular surface proteoglycans or was present in endosomes. We addressed this issue in two ways. First, to emove peptide bound to the cell surface, we repeated the above experiment, but included a brief incubation with trypsin following peptide treatment to remove surface bound peptide (Fig. 3.9). While small amounts

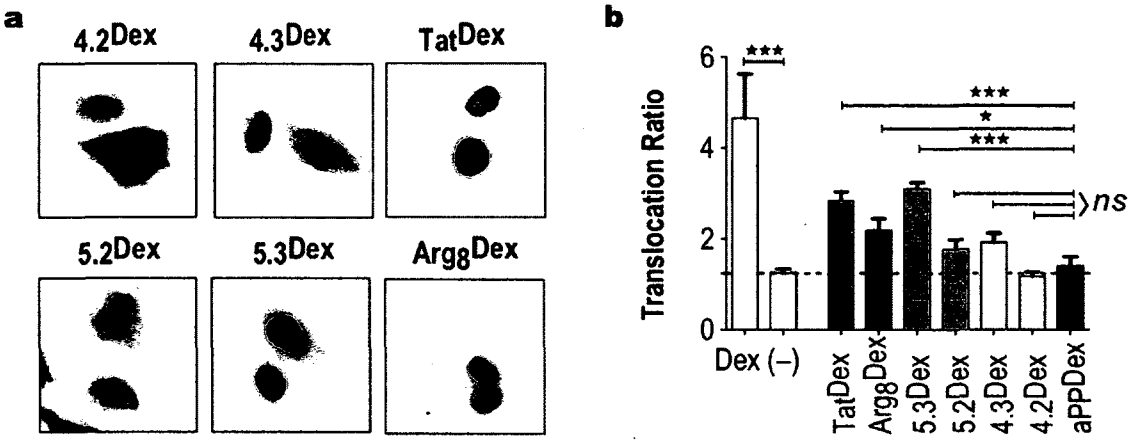


Figure 3.6: Translocation of GR-GFP after treatment with dexamethasone and dex-labeled miniature proteins. (**a**) HeLa cells transfected with GR-GFP (which appears black) after treatment with 1 μM 1 μM Dex-labeled cationic miniature protein for 30 min at 37°C. (**b**) Quantification of the changes visible in **a**. *ns*, not significant. $*p \leq 0.05$, $***p \leq 0.001$, ANOVA with Bonferroni post-test.

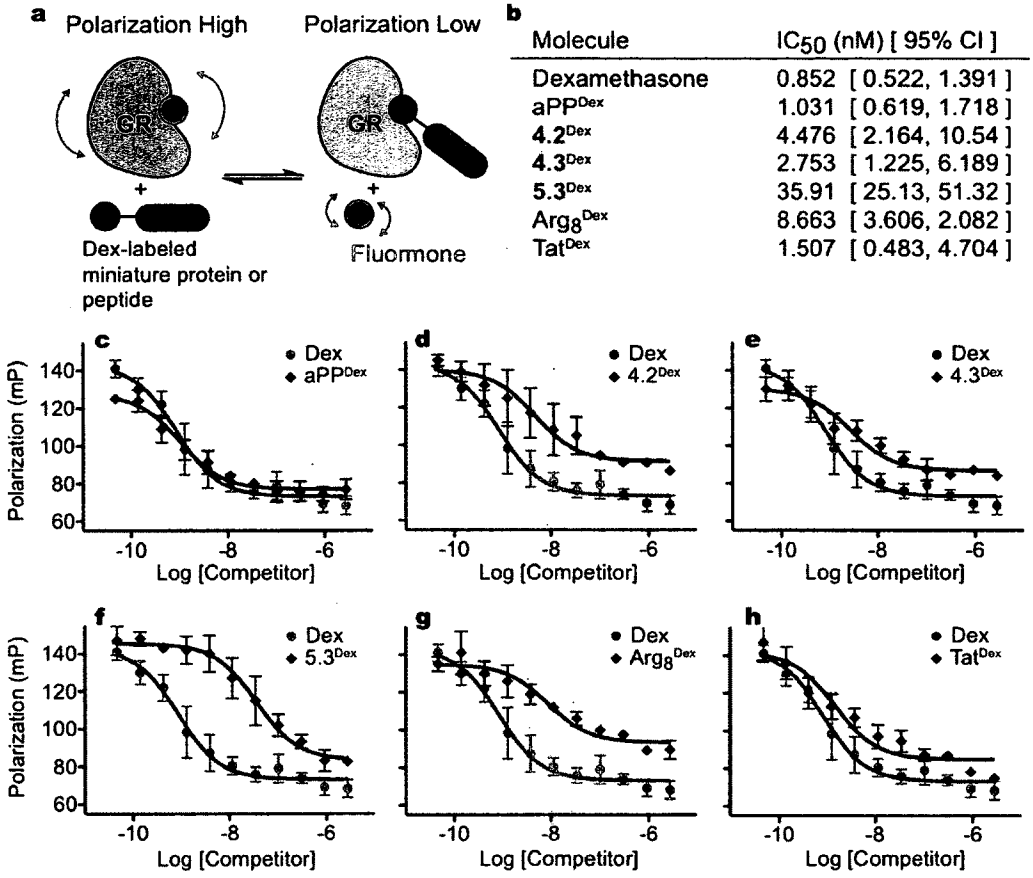


Figure 3.7: Competition binding of dexamethasone labeled cationic miniature proteins or peptides with GR. (**a**) Schematic representation of the assay; GR protein is mixed with Fluormone® (a fluorescently modified steroid) and varying concentrations of dexamethasone labeled miniature protein or peptide. Competition for the steroid binding site decreases the polarization of Fluormone. (**b**) Table of IC_{50} values and 95% confidence intervals (95% CI) obtained from fitting data in panels **c** – **h**. The data are fit to the equation $mP = P_{min} + (P_{max} - P_{min})/(1 + 10^{Log([Competitor])-Log\ IC_{50})})$. The ability of 5.3^{Dex} to compete with Fluormone® for the steroid binding site of GR is as good as or worse than that of dexamethasone or other Dex-labeled molecules tested. Competition binding of (**c**) $\mathbf{aPP}^{Dex}$, (**d**) 4.2^{Dex}, (**e**) 4.3^{Dex}, (**f**) 5.3^{Dex}, (**g**) $\mathbf{Tat}^{Dex}$, or (**h**) $\mathbf{Arg8}^{Dex}$ with Fluormone® for GR. Competition of dexamethasone with Fluormone® for GR binding is shown in grey for comparison. *[This data collected with Justin Holub.]*

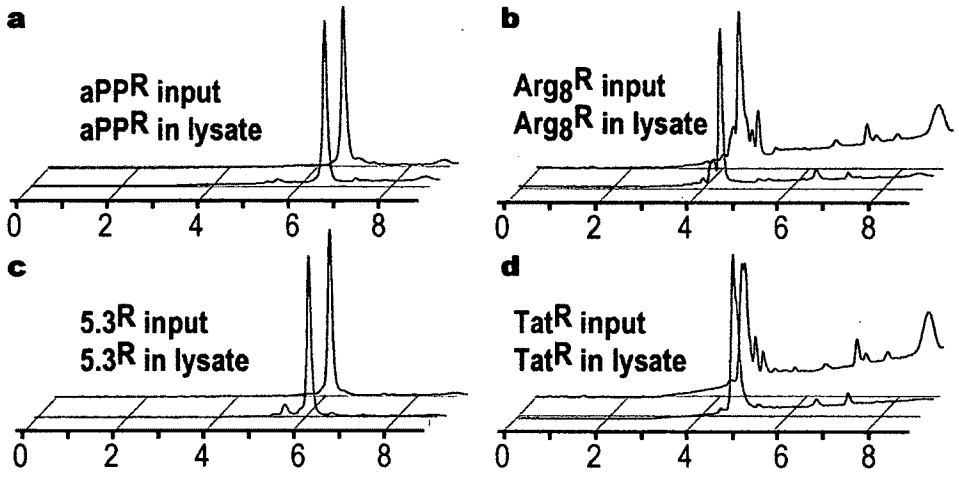


Figure 3.8: Absence of degradation of cationic miniature proteins. 2×10^5 HeLa cells were treated with 1 μM rhodamine labeled cationic miniature protein or peptide for 30 min in HKR buffer, before washing and lysis. Cell lysates were then examined by HPLC. HPLC traces of cationic miniature proteins are essentially unchanged between the cell lysate (black traces) and a control sample spiked with peptide not exposed to cells (red trace). Cell lysate from HeLa cells treated with rhodamine labeled **Arg8** or **Tat** show the appearance of numerous additional peaks, indicating modification or degradation of the peptide backbone.

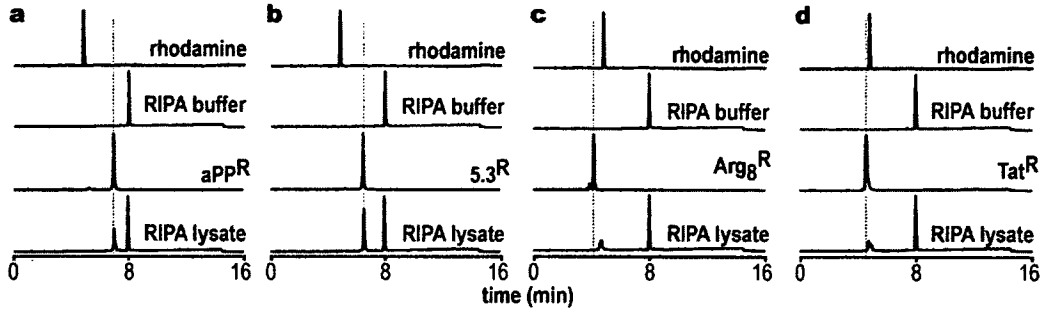


Figure 3.9: Absence of intracellular degradation of cationic miniature proteins. 2×10^5 HeLa cells were treated with 1 μM rhodamine labeled cationic miniature protein or peptide for 30 min in HKR buffer, washed twice and treated with 0.05% trypsin for 10 min at 37°C, after which the cells were washed again and lysed in RIPA buffer. Cell lysates were then examined by HPLC. HPLC traces of cationic miniature proteins are essentially unchanged between the cell lysate (red traces) and a control sample containing the starting material but not exposed to cells (blue trace). RIPA buffer alone shows a single peak (due to the presence of detergent, data not shown, grey traces). The retention time of rhodamine alone is shown as a control (black trace).

of **aPP**^{R} were recovered from cells after lysis and treated with trypsin (Fig. 3.9a), this amount was increased for cells treated with **5.3**^{R} (Fig 3.9b). In neither case was material with retention times distinct from starting material recovered. By contrast, multiple peaks were detected in lysates from cells treated with **Arg8**^{R} and **Tat**^{R} (Fig. 3.9c, d). These data confirm the results above, and suggest that increased recovery from cell lysates is due to the intracellular presence of **5.3**. Further more, the miniature proteins studied are minimally degraded within the time frame of the experiment.

Second, to specifically ask whether cationic miniature protein **5.3**, **Tat**, or **Arg8**, were present in the cytoplasm, we employed a cell fractionation strategy that exploits the bacterial pore-forming toxin streptolysin O (SLO) [146]. This peptide inserts into cholesterol containing membranes undergoing a conformational change[147] leading to oligomerization at 37°C creating membrane lesions that are approximately 20 nm in size[148] and are repaired in a Ca++ dependent manner[149]. At 37°C, and in the absence of Ca++, the presence of these membrane lesions leads allows soluble cytoplasmic contents to diffuse into the surrounding media (or surrounding proteins to diffuse into the cell[150]), while

larger organelles, and vesicles remain inside the cell. Thus incubation in the presence of SLO allows one to recover the cytoplasmic contents of a cell without disrupting endosomal membranes or releasing lysosomal contents. Others have cautioned, however, that even commercial preparations of SLO possess widely differing activities[151]. Incubation of cells with excess peptide can lead to aggregation and pore formation independently of cholesterol, decreasing the specificity of the process[151].

To verify that SLO induces cell permeabilization and to identify the concentration required for efficient recovery of cytoplasmic material, we incubated HeLa cells with various concentrations of SLO on ice using a protocol previously described[152, 153, 154, 146]. In brief, HeLa cells were detached from the growth flask by incubation in PBS containing EGTA/EDTA, washed and aliquoted into 96-well plates. The cells were then resuspended in ice cold solutions containing 5 mM dithiothreitol (DTT) in PBS with various concentrations of SLO added and incubated on ice for 15 min to allow the peptide to bind to the cell surface. (The presence of DTT is required for reduction of a critical cysteine within the SLO peptide[155].) The cells are then washed with PBS and shifted to 37°C to allow oligomerization of SLO inducing pore formation[147, 148] and releasing cytoplasmic contents. At this point, the cells can be pelleted and the supernatant containing the cytoplasmic extract recovered or the cells within the pellet tested further.

Following this protocol, HeLa cells were treated with PBS containing 5 mM DTT and between 7 μg/mL and 180 μg/mL SLO, leading to between 49% and 99% permeabilized cells (data not shown). HeLa cells treated with 20 μg/mL SLO showed between 85% and 89% +PI cells (Fig. 3.10a), indicating that this concentration of SLO permeabilized the majority of cells, that decreasing the amount of SLO would decrease the fraction of cells permeabilized. Because little advantage would be gained by increasing the amount of SLO, and due to the possibility of non-specific acitivity with higher concentrations, we tested this concentration for the release of the lysosomal enzyme β-hexosaminidase. HeLa cell treated with PBS or PBS containing SLO as above were pelleted by centrifu-

gation and the overlying supernatant recovered, while the cell pellet was then solubilized using a detergent (RIPA) buffer, which is known to release β-hexosaminidase, and in which this enzyme is active. Equal cell equivalents of supernatant from PBS and SLO treated cells, as well as the cell pellet from SLO treated cells were diluted in an acidic assay buffer containing 7.5 mM 4-nitrophenyl-N-acetyl-β-D-glucosaminide, a chromogenic β-hexosaminidase substrate. Hydrolysis by β-hexosaminidase leads to release of 4-nitrophenol, which absorbs 405 nm light, enabeling spectrophotometric quantitation. After incubation at 37°C for 1 hr, the reaction is stopped and the absorbance measured. The vast majority (>95%) of the enzymatic activity in SLO treated cells remained within the cell pellet (Fig. 3.10b), while only 2.5 ±0.1% of the β-hexosaminidase activity was present in the supernatant. These data indicate little, if any, permeabilization of the lysosomal membranes takes place when HeLa cells were treated under these conditions. Because β-hexosaminidase is present in early endosomes as well[156], it is unlikely that signficiant permeabilization of this compartment occurs. These data do not exclude the possibility that permeabilization of the secretory organelles such as the endoplasmic reticulum or the golgi apparatus occurs, however, using fluorescence microscopy we do not observe fluorescence in these compartments (see Figs. 4.6 on page 109, 4.1 on page 99, and 4.5 on page 108).

Having established that our SLO treatment protocol permeabilizes approximately 85% of the treated cells without disrupting endolysosomal membranes, we next asked whether we could detect intact or degraded rhodamine labeled cationic miniature proteins within cytoplasmic extracts. 2×10^5 HeLa cells were treated with 1 μM **aPP**^{R}, **5.3**^{R}, **5.2**^{R}, **Tat**^{R}, or **Arg8**^{R} at 37°C for 30 min before washing and tyspsin treatment. Cyotplasmic extracts were generated using SLO as described above and examined by fluorescence HPLC. Little or no fluorescence was detectable in cytoplasmic extracts treated with **aPP**^{R}, while fluorescence peaks were identified in cytplasmic extracts generated from cells treated with **5.3**^{R}, **5.2**^{R}, **Tat**^{R}, or **Arg8**^{R} (Fig. 3.11). Comparing starting material to recovered cytoplasmic extracts revealed that fluorescence material present in cytoplas-

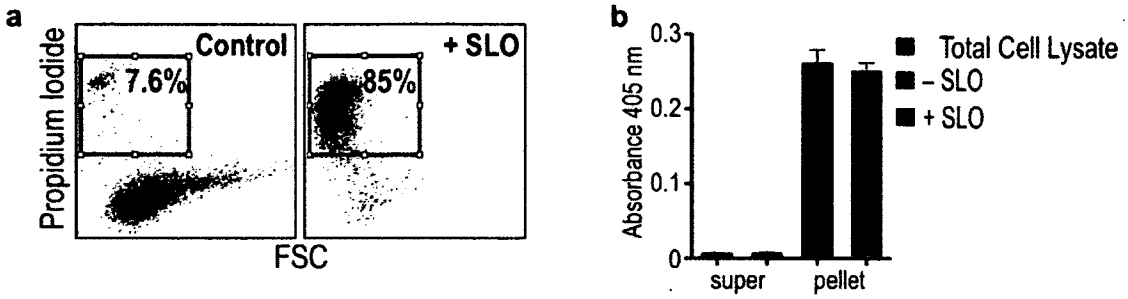


Figure 3.10: Confirmation that SLO induces cell permeabilization. (**a**) HeLa cells were treated with PBS (control) or PBS containing 20μg/mL SLO (+SLO) for 10 min on ice, after which they were washed and resuspended in PBS containing 1μg/mL propidium iodide. Exclusion of propidium iodide requires an intact cell membrane. Treatment with SLO increased the fraction of PI positive cells from 7.6% to 85%. (**b**) SLO treatment does not release the lysosomal enzyme β-hexosaminidase. HeLa cell treated with PBS or PBS containing SLO as in **a** were pelleted by centrifugation and the overlying supernatant recovered. The cell pellet was then solubilized using RIPA buffer. Equal cell equivalents of supernatant from PBS and SLO treated cells, as well as the cell pellet from SLO treated cells were incubated with a chromogenic β-hexosaminidase substrate. Enzymatic reactivity was assayed by increased absorbance at 405 nm. Greater than 90% of the enzymatic activity in SLO treated cells remains within the cell pellet.

mic extracts from cells treated with **Tat**^{R} (Fig. 3.11d) or **Arg8**^{R} (Fig. 3.11e) did not have the same retention times as the starting peptides. This could indicate covalent modification of the peptide itself or strong non-covalent association of cellular protein, however this latter possibility is unlikely because HPLC conditions (0.1% TFA in ACN/H_2O) are strongly denaturing. Cytoplasmic extracts from HeLa cells treated with **5.3**^{R} or **5.2**^{R} each showed single peaks of fluorescence material with retention times that matched starting material (Fig. 3.11b,c). The area under the peak was larger for **5.3**^{R} than **5.2**^{R}, consistent with the GR-GFP translocation assay (Fig. 3.6 on page 53) showing that greater nuclear translocation is observed after treatment with **5.3**^{R} than with **5.2**^{R}.

After endocytosis, internalized material is exposed to proteolytic hydrolases within the endolysosomal compartment, typified by the cathepsin family of acid-activated cysteine proteases. We wondered whether the enhanced activity of **5.3**^{R} within the GR-GFP translocation assay might be due to differential resistance or sensistivity to these pro-

teases. The former could increase the effective concentration within the endosome, the latter could release dexamethasone which is cell permeable. To distinguish between these possbilities, we treated **aPP**^{R}, **5.3**^{R}, **5.2**^{R}, **Tat**^{R}, and **Arg8**^{R} with endosomal cathepsins B, D, and L for 30 min at 37°C in acidic (0.1 M sodium citrate, pH 4.7) and reducing (5 mM DTT) conditions. The reactions were stopped by cooling to 4°C and dilution in 0.1% TFA in ACN/H_2O, prior to analysis by fluorescence HPLC (Fig. 3.11). Incubation with Cathepsin D completely modified the cationic miniature proteins test as indicated by the absene of any peak with the same retension times as the starting materials (Figs. 3.11a,b,c), while little difference was observed following incubation of Cathepsin D with **Tat**^{R} or **Arg8**^{R} (Figs. 3.11d,e). Incubation of **5.2**^{R} with cathepsin B resulted in the generation of multiple additional new peaks on the HPLC trace, although at the end of the 30 min incubation signficiant starting material remained indicating partial modification of this peptide (Fig. 3.11c). Similar results were obtained after incubation of **5.3**^{R} with cathepsin B, though the consumption of starting material was less extenstive (Fig. 3.11b). Cathepsin L failed to modifiy any of the cationic miniature proteins tested, but did completely consume **Arg8**^{R} (Fig. 3.11e). No detectable modification of **Tat**^{R} took place after incubation with cathepsins B, D, or L (Fig. 3.11d). These results indicate that digestion products of **aPP**^{R}, **5.3**^{R}, **5.2**^{R}, **Tat**^{R}, and **Arg8**^{R} by endosomal cathepsins are readily detectable by HPLC, suggesting that were these present within cytoplasmic extracts they would have been identified. Furthermore, because all cationic miniature proteins are susceptable to cathepsin B, these results decrease the possibility that the results observed are due to the differential sensitivities of these molecules to endosomal degradation. Were the results of the GR-GFP translocation assay due to release of dexamethasone following, for example, degradation of **5.2**^{Dex} by cathepsin D, we would expect this molecule to show more, rather than less activity than **5.3**^{Dex}.

These results suggest that quantitative comparisons of the amounts of cytoplasmic rhodamine labeled peptide should be possible by generating cytoplasmic extracts using SLO and comparing the amount of recovered peptide by integrating the area under the

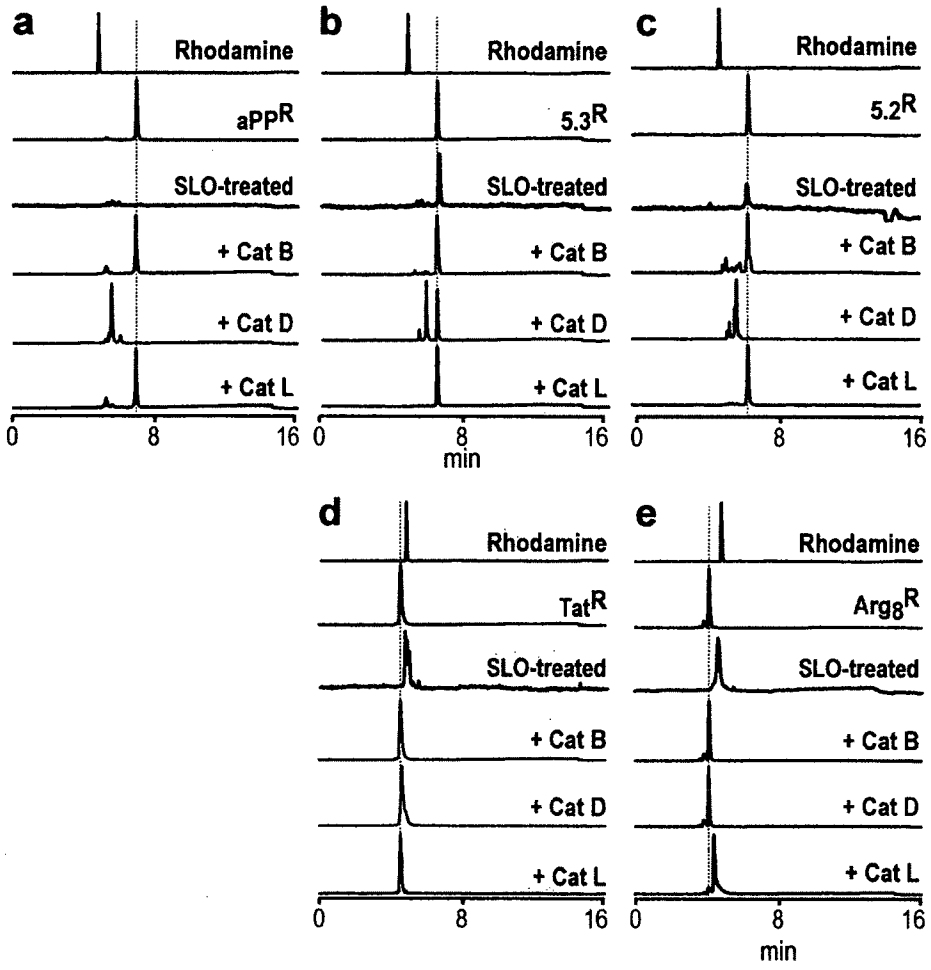


Figure 3.11: Recovery of cationic miniature protein 5.3, unmodified, from the cytoplasm. HPLC analysis of cytplasmic extracts generated using SLO from 2×10^5 HeLa cells treated for 30 min at 37°C with a solution containing 1 μM of the indicated peptides (green traces) are shown. HPLC alanysis of the purified peptides (starting material, red traces) and rhodamine alone (black traces) are shown for comparison. In addition, reactions containing 50 pmol peptide and 5pg of the indicated cathepsin were incubated at 37°C for 30 min prior to HPLC analysis (blue, purple, and green traces).

peak of the fluorescence HPLC trace. Data above support the hypothesis that these extracts are not significantly contaminiated by peptide present in the lysosome or bound to the cell surface. To quantitatively compare the relative fraction of $\mathbf{5.3}^R$ and $\mathbf{5.2}^R$ released into the cytoplasm, we treated duplicate samples with 1μM rhodamine labeled cationic miniature proteins $\mathbf{5.3}^R$ or $\mathbf{5.2}^R$ for 30 min at 37°C before either generating cytoplasmic extracts using SLO or total cell lysate using RIPA buffer (Fig. 3.12). This process was repeated in triplicate. Fluorescence HPLC traces for samples generated are shown in figures 3.12b and 3.12c. Quantitative integration of the identified peaks (Fig. 3.12d) in total cell lysates showed that for cells treated with $\mathbf{5.2}^R$ or $\mathbf{5.3}^R$, 11.7 ± 0.8 pmol and 10.1 ± 0.3 pmol per 2×10^5 HeLa cells were recoved, respectively. Cytoplasmic extracts of HeLa cells treated with $\mathbf{5.2}^R$ or $\mathbf{5.3}^R$ contained 0.31 ± 0.06 pmol and 0.44 ± 0.04 pmol per 2×10^5 HeLa cells, respectively. The fraction of $\mathbf{5.2}^R$ and $\mathbf{5.3}^R$ present in the cytoplasmic extracts (as compared to the total cell lysate) are $2.64 \pm 0.54\%$ and $4.37 \pm 0.32\%$ respectively, nearly 66% more for $\mathbf{5.3}^R$ than $\mathbf{5.2}^R$ ($p = 0.0516$). The variablility in these measurements makes quantitative *comparisons* preliminary, as the difference, though showing a strong trend, is not statistically significant.

This indicates that peptides $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$ are degraded or otherwise altered by the cells within the time course of these experiments. On the other hand, $\mathbf{5.3}^R$ remains unchanged.

Taken together these results support the conclusion that treatment with $\mathbf{5.3}^{Dex}$ induces the activation and nuclear translocation of GR-GFP because this peptide possesses an array of arginine residues that serve to target it to the cytoplasm.

3.2.6 Cationic Miniature Protein 5.3 is a monomer

Avian pancreatic polypeptide dimerizes[157, 158] in solution with a dissociation constant $K_D = 4.1 \times 10^{-6}\,\mathrm{M}^{-1}$ at pH 7[130]. It is possible that the incorporation of arginine residues leads to protein that is poorly behaved in solution. Aggregated protein can be

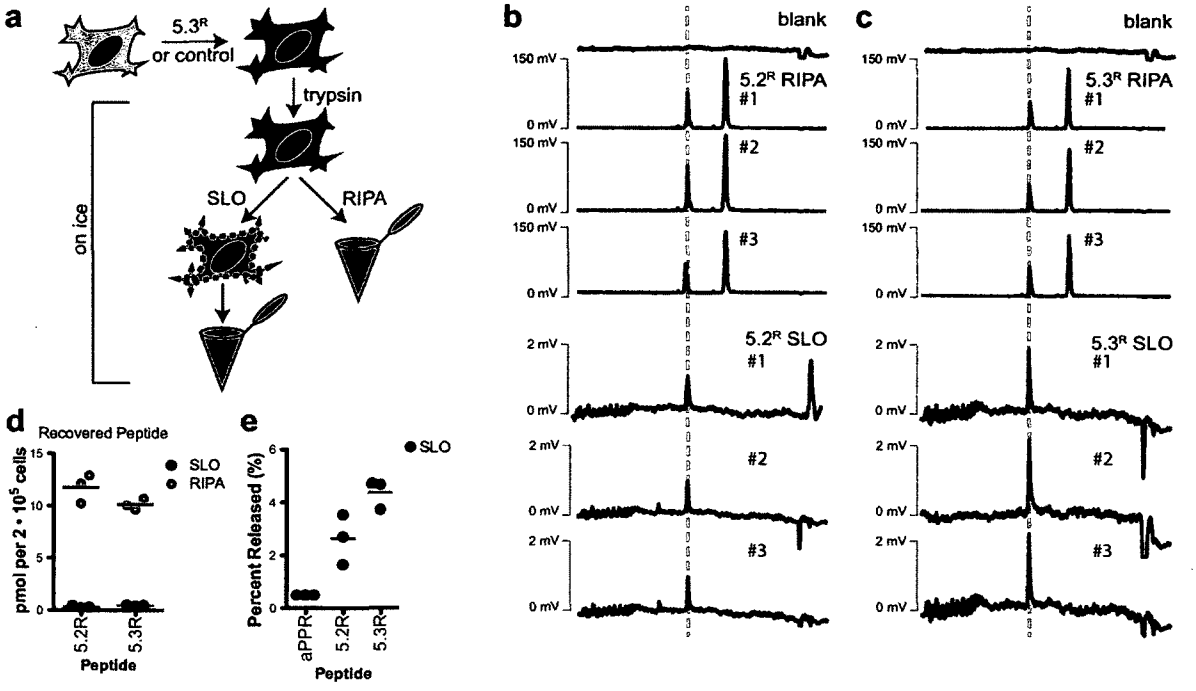


Figure 3.12: Direct observation that the 5.3 motif increases the cytoplasmic presence of miniature proteins. (**a**) Schema. Duplicate samples of 2×10^5 HeLa cells were treated with 1μM rhodamine labeled miniature protein for 30 min at 37°C. The cells were washed with 150μL PBS, then treated with 0.05% trypsin for 10 min at 37°C, and washed a second time. To generate total cell lysates, the sample was lysed by the addition of 40μL RIPA buffer and incubation on ice for 30 min. To generate cytoplasmic extracts, the cells were incubated with 50μL of an ice cold solution containing 20μg/mL SLO and 5 mM DTT in PBS for 15min at 4°C. The cells were washed twice with 150μL of PBS before warming to 37°C to allow pore formation. The cell pellet was separated from the cytoplasmic extract by centrifugation (2 min × 300*g*). HPLC traces of RIPA and SLO extracts from HeLa cells treated with (**b**) 5.2^R or (**c**) 5.3^R. The retention time of the starting peptide is shown by the vertical dashed grey line. The latter peak in the RIPA samples is due to light scatter from the presence of detergent. (**d**) Quantitative analysis of recovered peptide. The total amount of peptide recovered in RIPA and SLO extracts from samples is shown, based on standard curves generated from injecting known quantities of rhodamine and integrating the area under the major peak. (**e**) The fraction of peptide released into the cytoplasm, i.e. the fraction of peptide in the SLO sample compared to the RIPA sample, is shown for each of three replicates.

toxic to cells, and the aggregation of arginine rich peptides on the surface of cells can cause membrane disruption[126, 125]. While I previously showed that treatment of HeLa cells with 1 μM 5.3^{Dex} did not cause membrane disruption, I asked whether, at similar concentrations, 5.3 existed in a different associated state than other miniature proteins. I also wished to draw a comparison between the association state of unlabeled 5.3 and when labeled with dexamethasone, to understand whether installing the dexamethasone label altered the association state.

To accomplish this task, we turned to analytical ultracentrifugation (AU)[159], a well-established method to determine the association state of proteins[160], including aPP[157]. In this method, sample solutions containing protein are placed under a centrifugal force. After equilibrium is reached the concentration of protein as a function of the rotational radius (and thus angular velocity, ω) is determined using a spectrophotometer. The equilibrium profile obtained (usually plotted as protein concentration vs radius) depends on relative buoyant densities of the protein ($\overline{v}$) and solution (ρ), as well as the diffusion rate of the protein (D), which is sensitive to molecular size. The diffusion rate and the buoyant density are sensitive to oligomerization state. Obtaining sedimentation profiles under several different angular velocities allows fits to models using available software[161, 162].

The simplest model of sedimentation begins with a single ideal species, whose concentration at radius r from the center of rotation defines a sedimentation profile dependent on $\overline{v}$, ρ, ω, and molecular weight (M) as follows:

$$C(r) = C(r_0)e^{M(1-\overline{v}\rho)\frac{\omega(r^2-r_0^2)}{RT}}$$

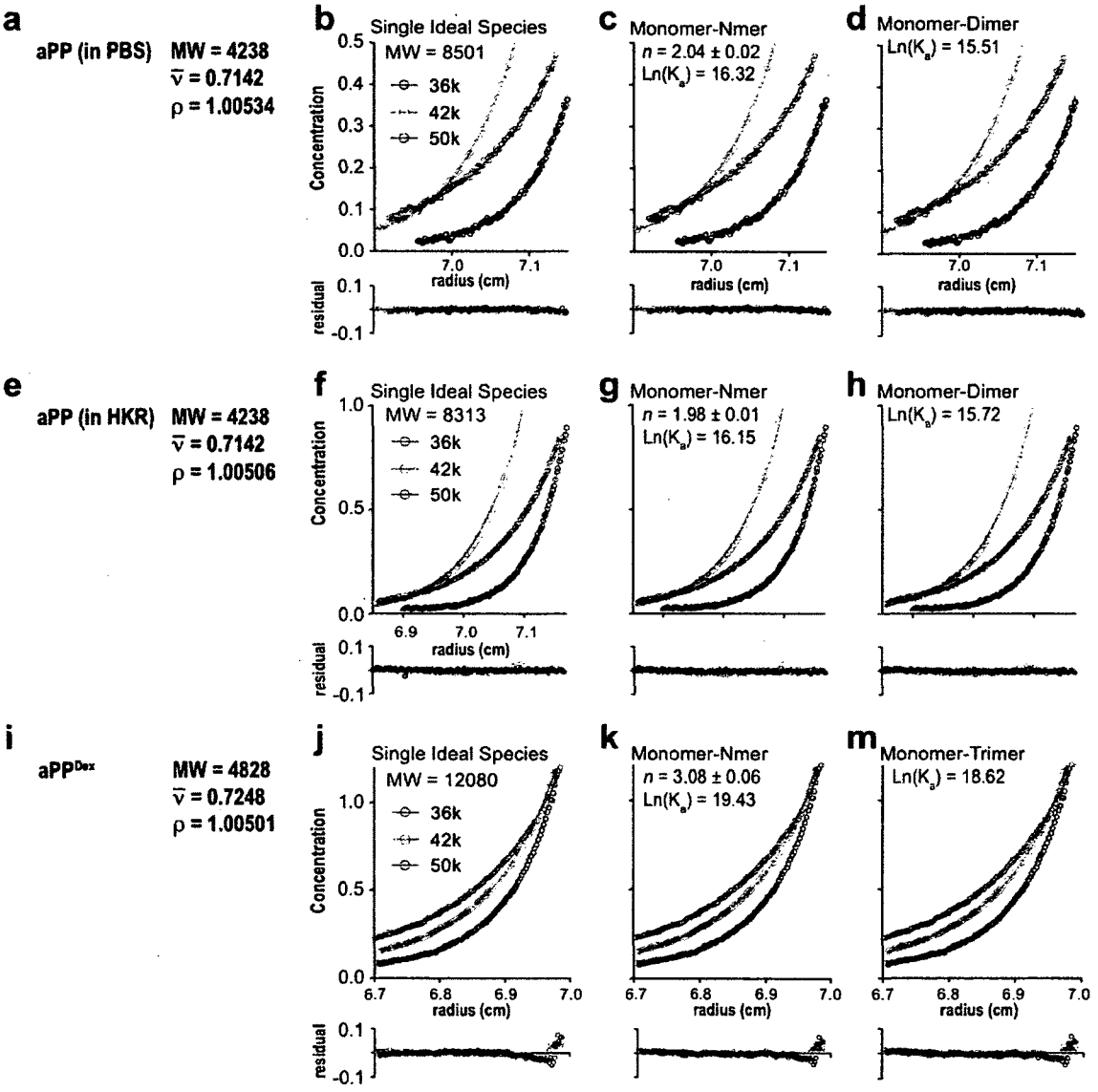
A general monomer-nmer self-associating system ($n \cdot A \rightleftharpoons A_n$) results in a superposition of two sedimenting ideal species (one with the molecular weight of the monomer, the other with a molecular weight n times larger), whose concentrations at all points along the solution column satisfy the law of mass action, and thus define an association

constant (K_a).

The examination of **aPP** by analytical ultracentrifugation has been previously reported[130, 157, 158]. Therefore we began by repeating the equilibrium sedimentation (Fig. 3.13) in the previously reported buffer (phosphate buffered saline, PBS), as well in the buffer used for the above studies of protein uptake (HKR buffer). Analysis of **aPP** in PBS showed self-association (fig. 3.13b-d). The single ideal species whose sedmidentation best-fit the data has a molecular weight of 8501 (which is 2.01 times the actual molecular weight of **aPP**, fig. 3.13b). Fitting the equilibrium sedimentation profiles to a monomer-dimer equilibrium (fig. 3.13d) yields $\ln K_a = 15.608$, which corresponds to a dissociation constant, $K_D = 1.66 \times 10^{-7}$ M. Repeating this analysis with **aPP** dissolved in HKR buffer, resulted in a sedimentation profile that also showed self-association. The molecular weight of a single ideal species that best fit the sedimentation profile was 8313 (1.96 times the molecular weight of **aPP**, fig. 3.13f). Fitting the sedimentation profile to a monomer-dimer equilibrium (fig. 3.13h) yielded $\ln K_a = 15.726$, or $K_D = 1.48 \times 10^{-7}$ M. These values are within the range previously reported ($K_D = 5 \times 10^{-6}$ to 2×10^{-8} M) [158, 130]. To test whether the installation of the dexamethasone label leads to a change in the association state, we repeated the above analysis using $\mathbf{aPP}^{Dex}$. This molecule also showed self-association (best-fit single ideal species MW = 12080, fig. 3.13j), however fitting the sedimentation profile to a general monomer-'Nmer' model revealed that the association was less energetically favorable. The model showing the best fit to the sedimentation

---

Figure 3.13 *(following page)*: Analysis of **aPP** and its dexamethasone labeled variant by analytical ultracentrifugation. The buoyant density of **aPP** (**a** and **e**) and $\mathbf{aPP}^{Dex}$ (**i**) were calculated according to the method of [163], while the density of the solutions were calculated based on the buffer components using the program SEDNTERP. The sedimentation profiles of **aPP** in PBS (**b**, **c**, **d**) and HKR (**f**, **g**, **h**) buffers as well as $\mathbf{aPP}^{Dex}$ in HKR buffer (**j**, **k**, **m**) after reaching equilibrium at 36,000, 42,000 and 50,000 rpm are shown. Fits to single sedimenting species (**b**, **f**, **j**) and general monomer-oligomer (**c**, **g**, **k**) are shown. For each general monomer-oligomer model, the sedimentation profiles were fit to an integer association state (**d**, **h**, **m**) closest to the best-fit variable n from the monomer-oligomer model. The residuals from each fit are shown below the graph of the sedimentation profile.


a
aPP (in PBS)
MW = 4238
$\bar{\nu}$ = 0.7142
ρ = 1.00534
b
Single Ideal Species
MW = 8501
36k
42k
50k
Concentration
radius (cm)
residual
c
Monomer-Nmer
n = 2.04 ± 0.02
Ln(K$_a$) = 16.32
d
Monomer-Dimer
Ln(K$_a$) = 15.51
e
aPP (in HKR)
MW = 4238
$\bar{\nu}$ = 0.7142
ρ = 1.00506
f
Single Ideal Species
MW = 8313
36k
42k
50k
g
Monomer-Nmer
n = 1.98 ± 0.01
Ln(K$_a$) = 16.15
h
Monomer-Dimer
Ln(K$_a$) = 15.72
i
aPP^{Dex}
MW = 4828
$\bar{\nu}$ = 0.7248
ρ = 1.00501
j
Single Ideal Species
MW = 12080
36k
42k
50k
k
Monomer-Nmer
n = 3.08 ± 0.06
Ln(K$_a$) = 19.43
m
Monomer-Trimer
Ln(K$_a$) = 18.62

profile was a monomer-trimer equilibrium, with $\ln K_a = 18.620$ (fig. 3.13m), which corresponds to a binding isotherm showing an inflection point (the concentration at which 50% of the miniature protein is monomeric) at 5.23×10^{-5} M. To aid in interpretation of these numbers we modeled the association states (see fig. 3.14) of **aPP** and **aPP**^{Dex} as a function of concentration, comparing results when using equilibrium constants previously reported[130]. The curves using the values for **aPP** in PBS and HKR buffers (determined in this study) overlap nearly completely (fig. 3.14a). The values determined previously show less extensive association of **aPP** at similar concentrations. The dexamethasone modified variant, **aPP**^{Dex} remains at least 99% monomeric at concentrations up to 5 μM. Thus, under the conditions used above, when **aPP**^{Dex} is added to cells at a concentration of 1 μM, the protein is monomeric.

We next set out to assess whether, when added to cells in the conditions used above, cationic miniature proteins **5.3** and **5.3**^{Dex} were also monomeric. Examination of **5.3** and **5.3**^{Dex} revealed that both of these miniature proteins weakly self-associate. Beginning with a model of a single ideal species and allowing the molecular weight to vary, the sedimentation profiles fit best to species of molecular weight significantly higher than the actual molecular weights of the miniature proteins. Using a general fitting strategy where the association state, n, and $\ln K_a$ are optimized yielded oligomerization states of 3.79 and 2.84 for **5.3** and **5.3**^{Dex} respectively. Fitting the sedimentation profiles to integer association states (i.e. 4 in the case of **5.3** and 3 in the case of **5.3**^{Dex}) yields a monomer-tetramer fit for the sedimentation profile of **5.3** with $\ln K_a = 31.995$, while a monomer-trimer fit for the sedimentation profile of **5.3**^{Dex} yields $\ln K_a = 21.029$. The concentrations at which 50% of the miniature proteins **5.3** and **5.3**^{Dex} are present as monomer are 15.7 μM and 17.9 μM, respectively. At the concentration used in the studies above (i.e. 1 μM), both the dexamethasone-labeled and unlabeled miniature protein are $>$ 99% monomeric. Thus there is little, if any, difference in the association state between **aPP**^{Dex} and **5.3**^{Dex} at the concentrations used. This data support the conclusion that the installation of arginines on the helical portion of **5.3** result in its uptake and cytoplasmic localization,

and that this effect is not due to change in the oligomerization state of the protein.

We found that at 1 μM all three miniature proteins (**aPP**, 5.3, and 5.3^{Dex}) are predominantly monomeric. Comparing the modeled transitions of 5.3 and 5.3^{Dex} results in nearly complete overlap, indicating that the weight fraction of monomeric miniature protein is not altered by the addition of the dexamethasone. Furthermore, at 1 μM, the concentrations at which these proteins are applied to cells, these miniature proteins are > 99% monomeric. Thus cationic miniature protein 5.3 is a monomer in solution at the concentrations used in these studies, and this conclusion is not dependent on the presence or absence of the dexamethasone label.

3.2.7 Natural Helices containing Arginine

The results reported here suggest molecules that contain a discrete α-helical arginine array can hijack endocytic machinery to cross into the cytosol. We wondered whether a similar α-helical arginine array might be found in natural proteins whose biological role demands the perturbation of membrane structure. Alternatively, if we found helical arginine arrays in proteins not known to interact with membranes, it would raise the possibility that membrane interaction was an unrecognized function.

3.2.7.1 An arginine-rich helix search

We began searching the PBD database for structures containing helices by extracting the amino acid sequence for each helix from the structure file. Structures deposited before 1995 were not analyzed, because these structures do not include records for secondary structure. My search yielded a list of 360,446 helical sequences from 37,479 peptide chains. Many of these represent reflected or homo- and hetero-oligomerized protein chains within the crystallographic asymmetric unit, and therefore share the same sequence as another helix within the same structure. After filtering out these occurrences, we are left with 217,015 helical sequences.

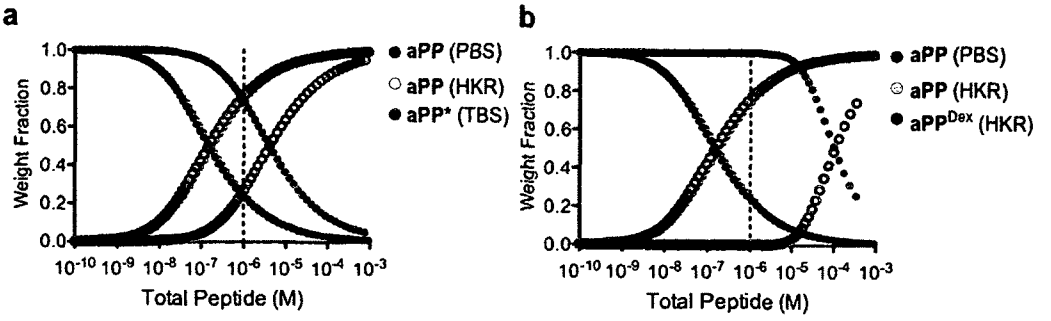
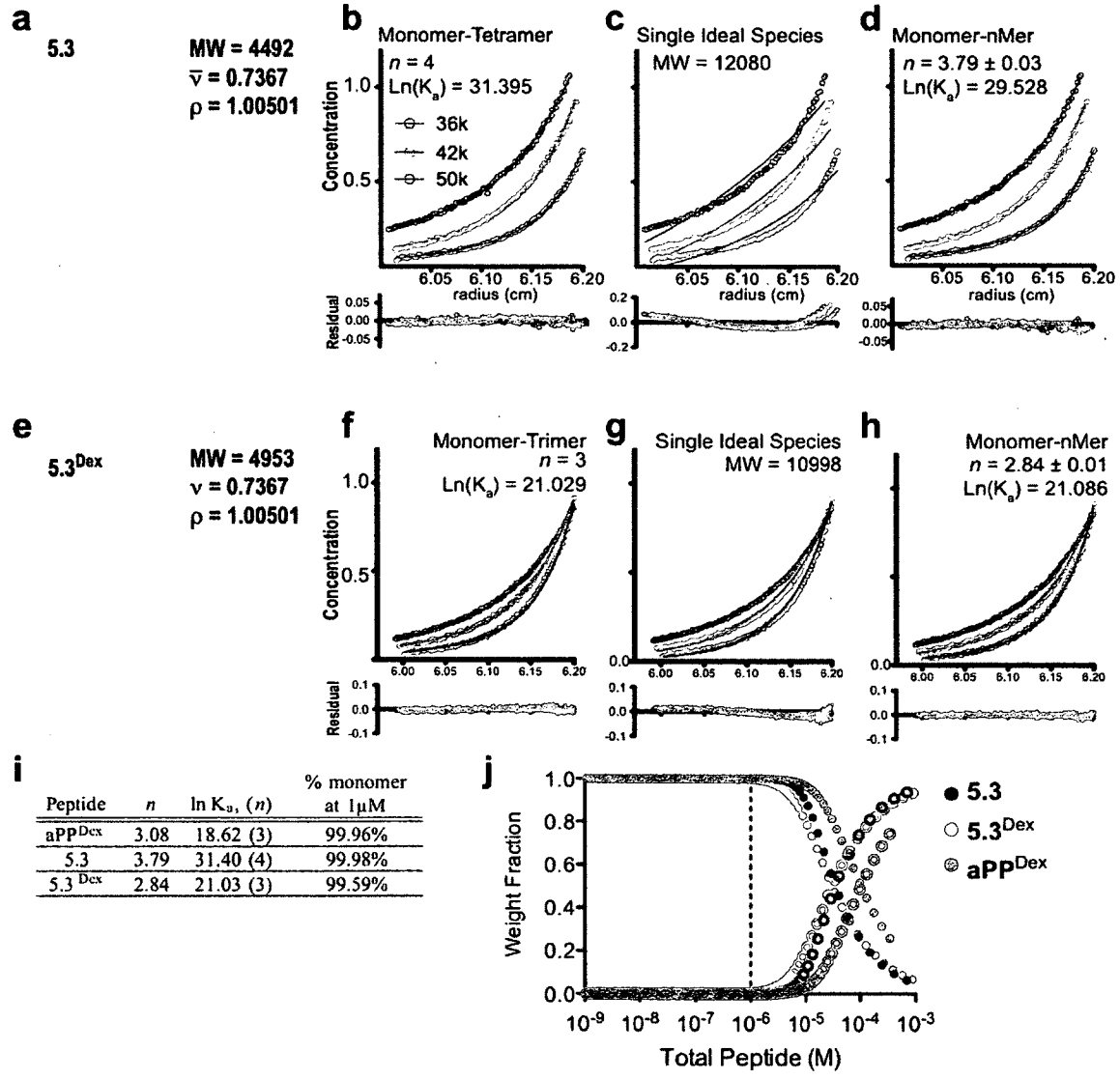


Figure 3.14: Association states of **aPP** and **aPP**^{Dex} as they vary with concentration. A self-consistency modeling approach was utilized to determine the weight fractions of monomeric (closed circles) and oligomeric (open circles) miniature protein by solving the mass action equation for self-associating systems at total protein concentrations between 10^{-10} and 10^{-3} M. The concentration used when studying the effects of peptides on cells (1 μM) is shown with a dotted green line. The parameters for **aPP** in either PBS or HKR buffer and **aPP**^{Dex} in HKR buffer were determined from the data shown in figure 3.13. The parameters for **aPP** in Tris-buffered saline (TBS; 25 mM Tris, 50 mM NaCl, pH 8.0) were reported previously[130]. (**a**) At similar concentrations, the fraction of oligomeric **aPP** is similar in PBS (blue) and HKR (orange) buffers, but lower in TBS. (**b**) Miniature protein **aPP**^{Dex} is monomeric at the concentration used in these studies. *[The Excel program for modeling the assembly of oligomers was contributed by Cody Craig.]*

Figure 3.15 *(following page)*: Cationic miniature protein **5.3** and its dexamethasone labeled variant are monomers. Physical parameters of **5.3** (**a**) and **5.3**^{Dex} (**e**) calculated by the method of [163]. The density of the solution, ρ, was calculated based on buffer components using SEDNTERP. The sedimentation profiles of **5.3** (open circles, **b**, **c**, **d**) and **5.3**^{Dex} (open circles, **f**, **g**, **h**) after reaching equilibrium at 36,000, 42,000, or 50,000 rpm are shown. (**b**) Best-fit monomer-tetramer model (lines) and residuals (below). The residuals do not show systematic deviation. The fit yeilds $\ln K_a = 31.395$. (**c**) Residuals from a model single ideal species shows systematic devation and a molecular weight higher than **5.3** indicating self-association. (**d**) A general monomer-nmer model fit reveals an association state $n = 3.79 \pm 0.03$. (**f**) Best-fit monomer-trimer model (lines) and residuals (below). The residuals do not show systematic deviation. The fit yeilds $\ln K_a = 21.029$. (**g**) Residuals from a model single ideal species shows systematic devation and a molecular weight higher than **5.3**^{Dex} indicating self-association. (**h**) A general monomer-nmer model fit to these data reveals an association state $n = 2.84 \pm 0.01$. (**i**) Summary table of the derived parameters from fitting sedimentation equilibrium profiles. For comparison, the values for **aPP**, reported by [130] are shown. (**j**)The weight-fraction of miniature protein existing as a monomer (closed circles) or in an oligomeric comples (open circles) in solution as a fuction of concentration. These data are modeled based on parameters shown in **i**. *[The Excel program modeling fraction monomers was contributed by Cody Craig.]*


a
5.3
MW = 4492
$\bar{\nu}$ = 0.7367
ρ = 1.00501
b
Monomer-Tetramer
n = 4
Ln(K$_a$) = 31.395
36k
42k
50k
Concentration
radius (cm)
Residual
c
Single Ideal Species
MW = 12080
d
Monomer-nMer
n = 3.79 ± 0.03
Ln(K$_a$) = 29.528
e
5.3^{Dex}
MW = 4953
ν = 0.7367
ρ = 1.00501
f
Monomer-Trimer
n = 3
Ln(K$_a$) = 21.029
g
Single Ideal Species
MW = 10998
h
Monomer-nMer
n = 2.84 ± 0.01
Ln(K$_a$) = 21.086
i
Peptide | n | ln K$_a$, (n) | % monomer at 1μM
aPP^{Dex} | 3.08 | 18.62 (3) | 99.96%
5.3 | 3.79 | 31.40 (4) | 99.98%
5.3^{Dex} | 2.84 | 21.03 (3) | 99.59%
j
Weight Fraction
Total Peptide (M)
5.3
5.3^{Dex}
aPP^{Dex}

To search for sequences enriched in arginine, we used a sliding window approach in which the number (and orientation) of arginines was tabulated within an 18 amino stretch, the length of the helix in 5.3. Of the 216,876 helical sequences, I found 2,531 that contained 4 or more arginines: 1649 sequences contained 4 arginines within an 18 amino acid window, while 399 contain 5 or more arginines. We next looked for helices in which the arginines were clustered in the same side of the helix. Over short stretches, the position of amino acids within α-helices show a seven-fold repeating arrangement over two turns of the helix, so that starting with position i on the first turn, $i+3$, $i+4$, are closely apposed on the rung above. Position $i+7$ is oriented above position i. Using the positional index of arginines and taking the numerical indices modulo 7, we filtered the arginine rich helices for those members that contain 4 or more arginines in the i, $i+3$, and $i+4$ position for any i along the helix length. We found 1484 helical sequences (derived from 449 structures) that fulfill these criteria. Because these sequences contain at least one set of three adjacent faces with a total of 4 arginine residues, they can be seen as variants of miniature proteins 4.3. This conception would place miniature proteins 4.2 and 5.2 as subsets of the 4.3 and 5.3 family. However, because these variants were less effective at obtaining cytosolic access, I chose to consider them separately. Filtering the above 449 structures for members containing exactly 4 or 5 arginine residues on three helical faces yields 14 protein chains containing 5 arginine residues within a window of 18 amino acids arranged on 3 helical faces. Interestingly, none of these proteins contained the exact arginine arrangement seen in 5.3 (arginines located in positions i, $i+4$, $i+7$, $i+10$, and $i+11$). A similar search for proteins containing exactly 4 arginines within an 18 amino acid window and located in i, $i+3$, and $i+4$ positions yielded 167 hits, identifying these proteins as members of the 4.3 family.

3.2.7.2 Functions of natural arginine rich helices

The list of 5.3 motif-containing natural proteins contains too few members to quantitatively confirm or deny the our hypothesis that this filtering algorithm enriches for

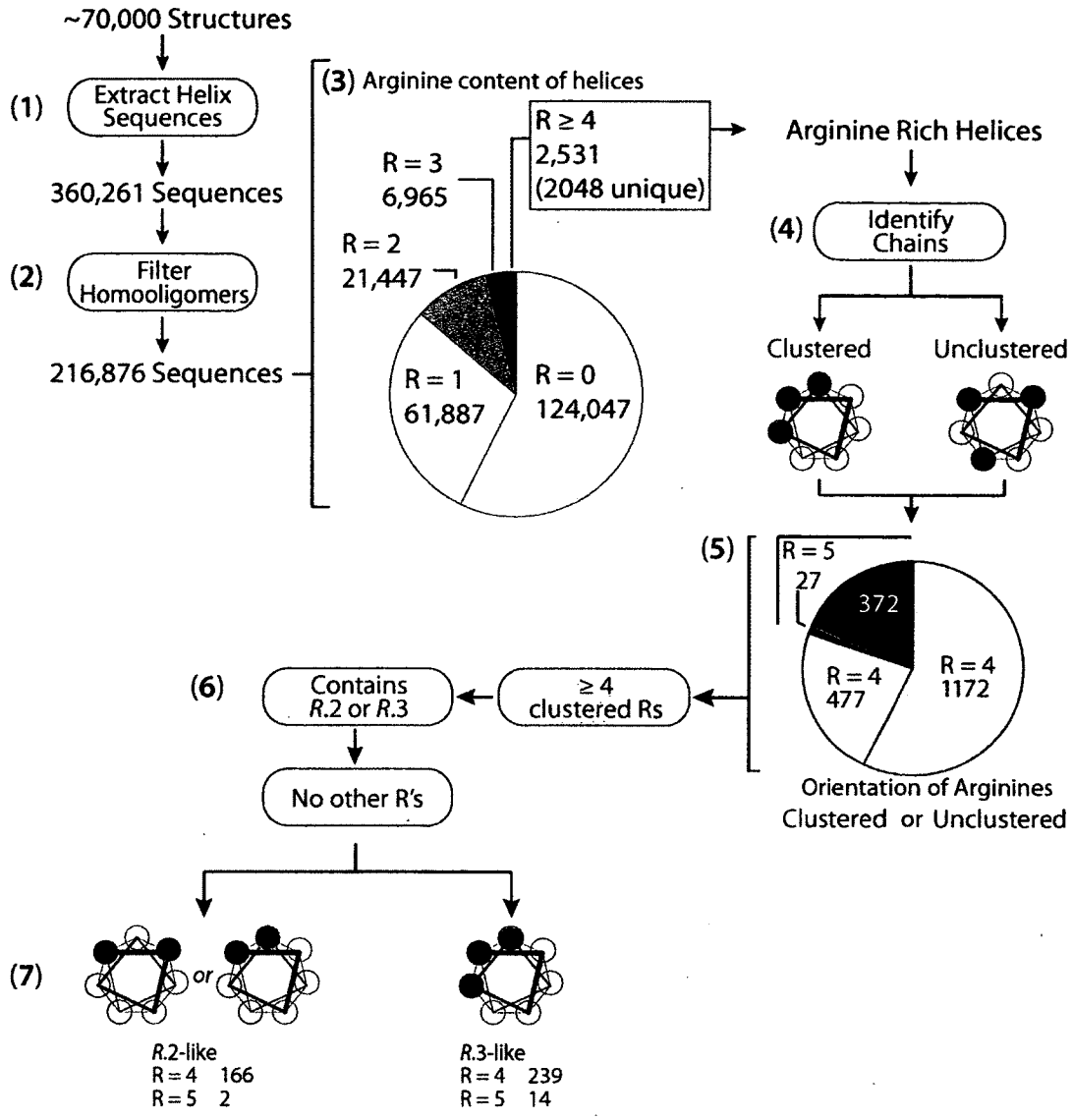


Figure 3.16: Schema for the identification of arginine-rich helices from the PDB. The entire PDB is searched for helical segments (1), which are then filtered for duplicated sequences within the same structure (2) and sorted based on arginine content (3). Out of 216,876 helical sequences, 2,048 unique sequences contain $\geq$4 arginine residues within an 18 amino acid window. Helical chains were then sorted as 'clustered' or 'unclustered', (4), a shown in the helical diagram. (5) 1172 sequences contained exactly 4 arginines, dispered across the entire helical surface. These were excluded from further analysis because they do not conform to the arrangement of arginines in 4.3. (6) Sequences containing $\geq$4 clustered arginines aranged on exactly two faces were designated '*R*.2-like', whereas sequences containing arginines arranged on exactly 3 faces were designated '*R*.3-like', where *R* is the number of arginine residues in the helix. These were filtered to exclude mutliple sequences from the same structure. (7) Results of the helix search: 239 4.3-like proteins and 14 5.3-like proteins were identified.

membrane-perturbing proteins. Members include a transcription factor (the thyroid nuclear hormone receptor) and several ribosomal proteins. It is unsurprising that proteins whose biological role is the interaction with negatively charged nucleic acid include positively charged segments. None the less, it is intriguing to observe that this set also includes CHMP3[164], a member of the ESCRT-III complex involved in membrane scission whose α-helical arginine array (identified in our search) is essential for its interaction with cellular membranes and ESCRT-III function[165]. The presence of the adenomatous polyposis coli (APC) tumor suppressor gene in our hit list is also intriguing. This proteins interacts tightly with the cytoskeleton at adherens junctions through a complex that includes Cdc42 and Rac1, proteins that link the process of endocytosis to the actin cytoskeleton. A previous study identified the activity of Rac1 as required for the actin reorganization that precedes the uptake of **Tat** and **Arg8**.

Broadening the search to examine the list of 4.3 motif-containing natural proteins results in additional insight, and strengthens the hypothesis that arginine rich helices interact with the actin cytoskeleton and are active motifs for uptake. Proteins containing arginine rich segments that play a role in the endocytic machinery either directly or through interaction with actin include complexin, spectrin alpha chain, phosphatidylinositol-binding clathrin assembly protein (CALM, a homolog of AP180), vinculin, and PLEKHA4, a PI3P binding protein. Also identified was BipD: a self-chaperoning protein found on the needle tip of the type 3 secretion system.

Literature searches illustrate the functional involvement of regions in which acidic residues are found, there are few examples of studies mutating arginine residues specifically. In the case of BipD, the helices in which the acidic residues are found are predicted to undergo a conformational shift, during the activation of the 'needle' to translocate virulence factors into target cells[166]. In the case of vinculin, the basic residues surround a patch that interacts with phosphoinositides[167]. In the case of CALM, these are not required for binding to $PI(4,5)P_2$ [168, 169, 170], but comprise the c-terminus of the CALM proteins ANTH domain and are conserved between CALM and AP180. Four arginine

PDB CODE	UniProt Accession	UniProt Name	Arginine Arrangement	Sequence
5.3-like protiens			a b c d e f g	
--	--	Cationic Miniature Protein 5.3 (*this study*)	2 0 0 1 2 0 0	RDLIRFYRDLRRYLNVV
2GD5	Q9Y3E7	Charged multivesicular body protein 3 (CHMP3)	1 2 0 0 2 0 0	RKEXRVVDRQIRDIQREE
1EH1	Q9WX76	Ribosome-recycling factor	2 0 0 3 0 0 0	RAVRQYAEEGRVAIRNIR
1YSH	Q5I7K9	60S ribosomal protein L30	2 0 0 2 0 0 1	RLILVESRIHRLARYYKR
1WQ6	Q06455	Protein CBFA2/Fusion Protein AML-ETO	0 0 0 3 2 0 0	EKTRRSLTVLRRCQEADR
1MHE	P61769	HLA class I histocompatibility antigen, alpha chain E	1 0 0 2 0 0 2	RETRSARDTAQIFRVNLR
1M5I	P25054	Adenomatous polyposis coli protein	2 0 0 2 1 0 0	RAQRRIARIQQIEKDILR
3M9E	P18113	Thyroid hormone receptor beta	0 0 2 1 0 0 2	SKRLAKRKLIEENREKRR
1Y69	P0A807	Ribosome-recycling factor	3 1 0 0 1 0 0	RGEAEQARVAVRNVRRDA*
1EK8	P0A805	Ribosome-recycling factor	3 1 0 0 1 0 0	RGEAEQARVAVRNVRRDA*
1IS1	Q8GRF5	Ribosome-recycling factor	3 1 0 0 1 0 0	RGEAEGGRVAVRNIRRDA
2WWL	P0A7W7	30S ribosomal protein S7	1 0 0 3 1 0 0	RPVRRNALAMRWIVEAAR
1KEJ	P09838	DNA nucleotidylexotransferase	0 3 1 0 0 1 0	SRQFERDLRRYATHERK
1KA2	Q8U3L0	Carboxypeptidase 1, unreviewed	2 0 0 2 1 0 0	EYERGIVRVLDRSIRIAR
1QQN	P19120	Heat shock cognate 71 kDa protein (Hsc70)	0 0 2 0 0 1 2	NKRAVRRLATACERAKRT
4.3-like protiens (selected)				
1ID3	P61830	Histone H3	0 0 1 2 0 0 1	PVGRVHRLLRR
3NFT	Q63K37	Translocator protein BipD	1 0 0 2 0 0 1	TVFDDARVAVRGHARAQR
1TR2	P18206	Vinculin	1 0 0 2 0 0 1	GATSIARRANRILLVAKR
1UPQ	Q9H4M7	Pleckstrin homology domain-containing family A member 4	0 0 2 0 0 1 1	TLEDLRGWLRALGRASR
16VP	P06492	VP16	0 0 1 1 0 0 2	FCSALYRYLRASVRQLHR

Figure 3.17: Natural proteins containing arginine-rich helices identified from searching the PDB. Proteins containing helical segments that are consistent with the 5.3 and 4.3 motifs are shown. The PDB code for each protein, the corresponding Uniprot accession number, and the protein name are shown together with the number of arginine residues in each helical postion as well as the amino acid sequence of the arginine rich segment.

residues also present within the same c-terminal region of the ENTH domain but these residues are split over two-helices with a kink in the middle. The biological function and the identity of the lipid interacting residues of PLEKHA4 remain untested.

Whether or how these proteins utilize arginine rich helices remains unknown. Arginine rich helices may play a specific role in modifying membrane structure through interaction directly with lipid, or may serve other functions. Drawing functional comparisons between cationic miniature proteins and natural cytoplasmic proteins that contain arginine rich helices is complicated by the fact that these proteins exist (at least initially) on the opposite side of the cell membrane. Nonetheless, the presence of cytoplasmic proteins that contain this structural motif is encouraging. The role of BipD may be more analogous to cationic miniature proteins because it functions on the extracellular face of membranes. Experimental comparisons may reveal unknown functions of basic residues within the type 3 secretion system, alternatively studying the mechanisms of proteins containing arginine rich helices may reveal clues that improve our ability to design cytosolic miniature proteins.

Thus, compared to searches looking only for overall positive charge[119, 118], our analysis of structural determinants for miniature protein entry greatly expands the number of potential proteins in nature that may contain cell–penetrating or at least membrane–perturbing function. The ability of cationic miniature protein 5.3 to rapidly escape from early endosomes is a unique feature, not shared by canonical cell penetrating peptides, and represents a starting point for the optimization of well-folded functional cell penetrating proteins useful as pharmacologic tools capable of modulating cytoplasmic protein function. More sophisticated comparisons than the one presented here may illuminate shared features of these and other arginine rich proteins.

3.3 Discussion

The interfaces that form between and among proteins and DNA–often large, flat, and polar–do not resemble those that bind small molecule substrates or traditional inhibitors[132]. Targeting these 'undruggable' interfaces is a task well suited to protein and peptide ligands, but can only be successful if such molecules reach their cytosolic targets. Unfortunately, the very properties that endow peptide mimetics with their promise–size and polarity–are precisely those properties forbidden by Lipinski's rules[2]. The challenge, therefore, is to identify the determinants that guide the uptake of peptide-like molecules and the mechanisms through which they gain cytosolic access, generating a new set of rules applicable to large peptidic molecules. Advancing this goal has been constrained by the absence of a rapid and robust assay capable of distinguishing between peptide-like molecules trapped within endosomes or those that have escaped into cytosol.

Here we took advantage of previous observations that the cytoplasmic-to-nuclear translocation of the glucocorticoid receptor (GR) takes place within minutes[144] and recent developments in automated high–content image analysis, to develop an assay that exploits the nuclear to cytoplasmic ratio of a GR-GFP fusion protein as a proxy for the cytoplasmic entry of traditionally impermeant molecules tagged with dexamethasone. Compared with a first-generation assay developed by Kodadek and colleagues[143], the one reported here provides a readout in intact living cells within 30 min (as opposed to 48 hours) using an epifluorescent microscope and eliminates the requirement for cell lysis or enzymatic substrates, substantially reducing cost.

Beginning with a panel of miniature proteins that contained variety in the number and orientation of arginine residues, I found that the structure of the miniature protein scaffold **aPP** was tolerant to incorporation of up to five arginine residues. Next, by synthesizing fluorescently labeled variants, I found that the incorporation of arginine resulted in molecules that were taken up more efficiently by cells that the parent molecule **aPP**. For molecules containing four arginines, the magnitude of the effect dependend on

the relative orientation of the positively charged residues, with the clustered arginines present in 4.2 resulting in greater uptake than the more dispersed arginines found in 4.3. In the case of molecules containing five arginines, the magnitude of the effect was similar regardless of their orientation.

Among the four miniature proteins I identified as being actively taken up by cells, I found clear differences in the ability to activate the GR when covalently linked to dexamethasone. GR activation by dexamethasone-labeled peptides and miniature proteins implies their presence in the cytoplasm. An alternative hypothesis is that the peptides and miniature proteins studied are substantially degraded during the course of the experiment, leading to spurious activation of the GR by released (and perhaps cell permeable) dexamethasone-peptide fragments. HPLC analysis of aPP^{R} and 5.3^{R} recovered from cell lysates (Fig 3.8) shows that miniature proteins remain unmodified during the course of the experiment excluding this possibility. Thus, the activation of the GR is due to the arrival in the cytoplasm of the intact miniature proteins, and not due to the effect of a degradation product.

For many cell penetrating peptides, cellular uptake results from an endocytic process leaving the peptides trapped within a membrane bound endosome and unable to reach their cytosolic target. It is tempting to invoke a hypothesis that endosomes are 'leaky' and that peptides escape to the cytoplasm with similar rates. Even if this were this the case, the differences in GR activation by 4.2^{Dex}, 4.3^{Dex}, 5.2^{Dex} and 5.3^{Dex} are not explainable by differences in their endocytic rate (4.2^{R}, 5.2^{R}, and 5.3^{R} are each taken up to a similar extent). Indeed 4.3^{R} is taken up to a lesser extent than 4.2^{R}, but 4.3^{Dex} is a superior activator of GR than is 4.2^{Dex}, though the difference is small. Differences in affinity for the GR receptor itself also do not explain these differences, as the *in vitro* affinity of 4.2^{Dex} and 4.3^{Dex} for the GR is similar. A parallel argument holds true for peptides 4.2 and 5.3: both are taken up to a similar extent, but treatment with $1\mu M$ 5.3^{Dex} leads to greater activation than does treatment with 4.2^{Dex}. This comparison is strengthened by the fact that 5.3^{Dex} has a *decreased* affinity for the GR, requiring a *higher* effective concentration

to achieve the same fractional occupancy. Taken together, these findings identify 5.3 as a cationic miniature protein capable of efficiently reaching the cytosol without disrupting the cellular membrane.

How do the findings of 5.3 compare to cannonical peptides used to impart cytosolic access? This comparison is somewhat complicated by the fact that lysates from cells treated for 30 min with $\mathbf{Tat}^{R}$ or $\mathbf{Arg8}^{R}$ show the appearance of degradation products. This raises the possibility that the GR induction seen after treatment with $\mathbf{Tat}^{Dex}$ or $\mathbf{Arg8}^{Dex}$ may be, in part, due to the action of degradation products rather than the intact peptide. These data do not distinguish whether the eventual degradation of $\mathbf{Tat}^{R}$ and $\mathbf{Arg8}^{R}$ takes place within endosomes or the cytoplasm. If these molecules are degraded within the cytoplasm, they may remain useful as modules to direct peptidomimetics to cytoplasmic targets. The observed instability of **Tat** and **Arg8** stands in contrast to the stability of 5.3, a benefit of well-folded miniature proteins. This caveat not withstanding, direct comparisons of 5.3^{Dex} to $\mathbf{Tat}^{Dex}$ and $\mathbf{Arg8}^{Dex}$ reveal that translocation ratios are similar for 5.3^{Dex} and $\mathbf{Tat}^{Dex}$ ($p = 0.147$), but that the translocation ratio after treatment with 5.3^{Dex} is higher than after treatment with $\mathbf{Arg8}^{Dex}$ ($p = 0.3 \times 10^{-9}$) suggesting that **Arg8** is less efficient at reaching the cytoplasm.

Comparison among these three peptides further supports the conclusion that internalization and cytoplasmic access are separable, as cellular uptake of $\mathbf{Arg8}^{R}$ is superior to $\mathbf{Tat}^{R}$, both of which are inferior to 5.3^{R}. Thus, some of the peptides studied here contain arginine motifs that display only efficient uptake (**4.2**, **5.2**) or only cytoplasmic localization (**Tat**). While **Arg8** is intermediate in both uptake and cytoplasmic access, miniature protein 5.3 effectively combines both of these features and displays increased proteolytic stability. These comparisons reveal that the process of cellular uptake and cytoplasmic access are governed by discrete structural determinants and that peptides and miniature proteins may be optimized to possess either or both of these functions.

In summary, this study led to the identification of 5.3, a cationic miniature protein whose rapid cytosolic localization is equivalent to or better than **Tat** and **Arg8**. This

comparison revealed that contrasting structural determinants control cellular uptake and cytoplasmic localization. Uptake is favored by clustered α-helical arginine side chains, whereas cytoplasmic access requires a more dispersed arginine array.

How can the arginine residues mediate cytosolic localization? The differences in arginine/lysine number and orientation in **5.3**, **Tat**, and **Arg8** can affect side chain pK_a (and thus overall charge at a given pH), the ability to form neutral ion pairs with phospholipid head groups, or affinity for as-yet-unidentified cellular targets. How this difference in charge state and or distribution could result in increased cytosolic access is unknown. If these molecules are internalized into endocytic vesicles, it is possible that the distinct arginine array in **5.3** favors the formation of a critical protonated species in the moderately acidic environment of the early endosome ($pH \approx 6.5$). **Tat** and **Arg8** on the other hand, may require the lower pH present in lysosomes inorder to generate an equivalent. It is equally possible that the discrete arginine array in **5.3** represents an specific export signal for cellular machinery that has not yet been identified.

3.4 Experimental Methods

3.4.1 Miniature protein synthesis and purification.

All miniature proteins and peptides were synthesized using a Liberty Automated Microwave Peptide Synthesizer (CEM) and standard solid phase synthesis techniques. Coupling reactions were performed using 5 equivalents of amino acid, 5 equivalents HBTU, and 10 equivalents of diisopropylethylamine in DMF. FMOC groups were removed using 20% piperidine in DMF containing 0.1 M HOBT to minimize aspartamide formation2. Following synthesis, the peptides were cleaved from the resin by adding 5 mL of a solution containing 95% TFA, 2.5% water, 2.5% TIPS and incubating at 38°C for 30 min using a microwave accelerator (MARS system, CEM). The solution containing the crude peptide was precipitated by the addition of 40 mL cold ether. The precipitate was dried under a nitrogen stream and lyophilized before purification via reverse phase

HPLC (YMCbasic column, 150mm x 10mm ID) using a linear gradient of CH_3CN in water (both acidified with 1% TFA). Purified miniature proteins or peptides were collected and lyophilized, then reconstituted in water and kept at -4 °C protected from light. The purity and identity of purified miniature proteins was characterized by analytical HPLC (>95% in all cases) and either a Voyager (Applied Biosystems) MALDI-TOF (matrix-assisted laser desorption-ionization time-of-flight) spectrometer (337 nm laser, α-cyano-4-hydroxycinnaminic acid matrix), or a Waters QTof LC-MS. For molecular formulas, calculated masses, and measured (m/z) ratios of polypeptides and miniature protein domains see figure 3.20. Labeled miniature proteins and peptides were assessed for purity by analytical HPLC. Peptides and miniature proteins were injected on a Shimadzhu HPLC equipped with a fluorescence detector, a UV-Vis detector and fitted with an Agilent Poroshell 120 SB-C18 column (100 mm x 3.1 mm ID, 2.7 μm particle size). A linear gradient of 10 - 100% CH_3CN in water (both acidified with 0.1% TFA) was used as a solvent system. Fluorescein labeled compounds were detected via fluorescence (excitation 450 nm, emission 500 nm). Rhodamine labeled compounds were detected via fluorescence (excitation 550 nm, emission 580 nm). Dexamethasone labeled compounds were detected via absorbance at 280 nm. Traces are shown in figure 3.18 and figure 3.19.

3.4.1.1 Fluorescein labeled miniature proteins.

Fluorescein labeled miniature proteins were generated from purified peptides containing an appended C-terminal cysteine residue. The purified unlabeled product was reacted with 5-(iodoacetamido)fluorescein to generate the fluorescein-labeled miniature protein as follows: to 500 μL of a 100 mM phosphate buffer, pH 7.25 was added 1 μmol of the miniature protein, then water to a total volume of 2.5 mL. Separately, 5 mg of 5-(iodoacetamido)fluorescein (Sigma Cat. #I9271) was dissolved in 2.5 mL DMF. This DMF solution was added to the aqueous miniature protein solution and the reaction was incubated in the dark at 25 °C for 2 hours. To generate unlabeled miniature proteins, the C-terminal cysteine was capped by reaction with iodoacetamide under analogous con-

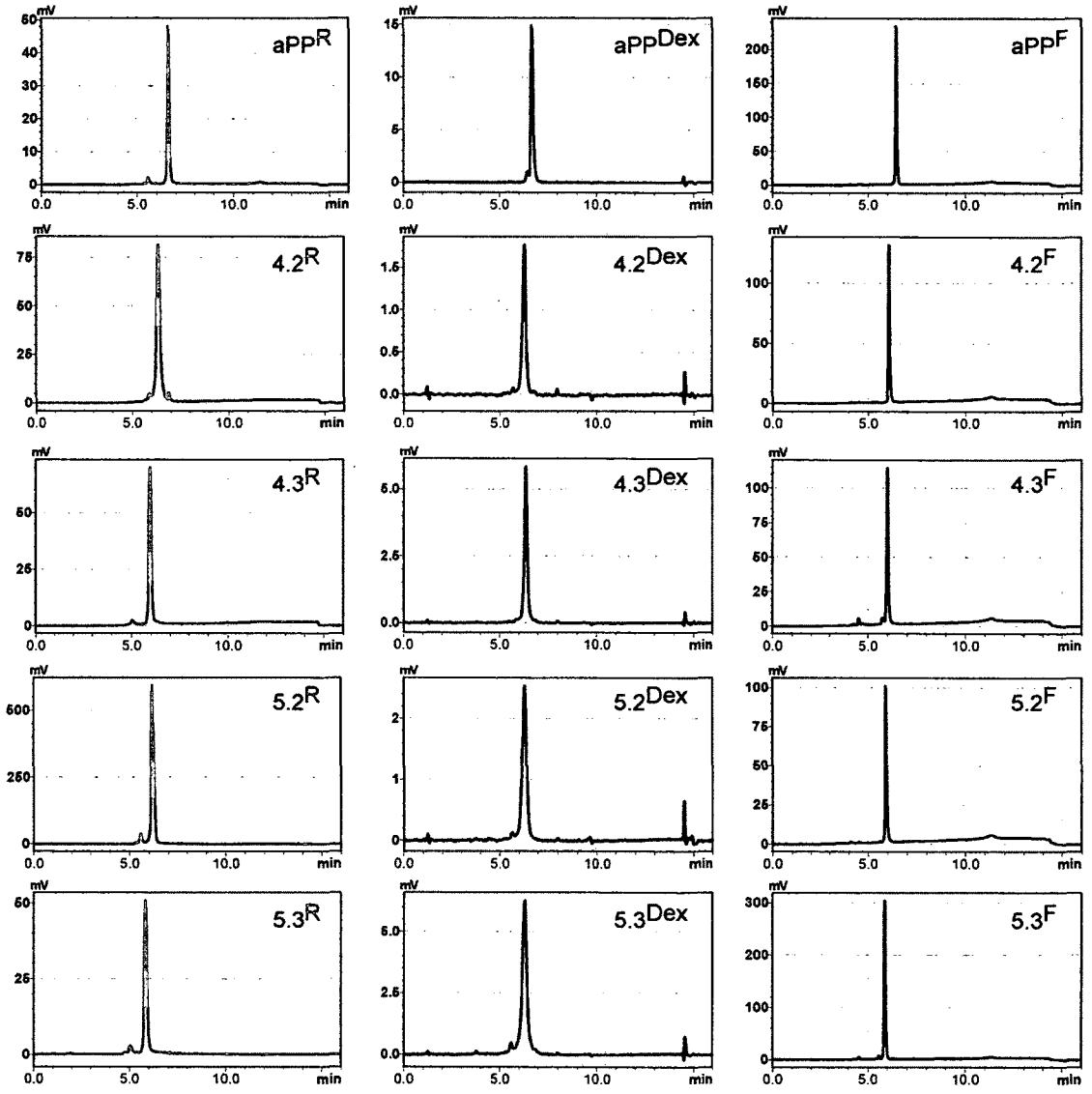


Figure 3.18: HPLC analysis of selected cell permeable miniature proteins.

ditions. The reaction mixture was purified using reverse phase HPLC using the column and solvent system described above.

3.4.1.2 Rhodamine labeled miniature proteins and peptides.

Rhodamine labeled miniature proteins were generated by carrying out labeling reactions on resin using an appended N-terminal lysine. For this purpose we employed an amino acid monomer in which N^{α} is BOC-protected and N^{ϵ} is protected with an FMOC group. The N^{ϵ} was deprotected using microwave acceleration and incubating the resin twice with 20% piperidine in DMF at 80°C for 7 min. The resin was then washed thoroughly with DMF. To label peptides with tetraethyl rhodamine, to the deprotected resin was added 10 eq of Lissamine Rhodamine B sulfonyl chloride and 10 eq of diisopropylethylamine in 5 mL DMF at room temperature overnight. The resin was thoroughly washed with DMF and DCM before global deprotection and cleavage using the conditions for unlabeled peptides described above. Crude rhodamine labeled peptide was precipitated in 40 mL of cold ether and lyophilized. Peptides were purified by reverse phase HPLC as described above. Peptides lyophilized from HPLC solvents were stored in water at 4°C protected from light.

3.4.1.3 SDEX labeled miniature proteins and peptides.

SDEX labeled miniature proteins and peptides were generated in a manner analogous the generation of rhodamine labeled miniature proteins; after appending an N-terminal lysine residue bearing an FMOC protecting group in the side chain, the N^{ϵ} was deprotected using microwave acceleration by incubating the resin twice with 20% piperidine in DMF at 80°C for 7 min. The resin was then washed thoroughly with DMF. To the deprotected resin was added 5 eq SDEX, 5 eq HATU, 5 eq of HOAT, 5 eq of DIPEA, and 5 eq of 2,6-lutidine in 5 mL DMF. The reaction was shaken overnight at room temperature. The resin was thoroughly washed with DMF and DCM before global deprotection and cleavage using the conditions for unlabeled peptides described above. Crude SDEX

labeled peptide was precipitated in 40 mL of cold ether and lyophilized. Peptides were purified by reverse phase HPLC as described above. Peptides lyophilized from HPLC solvents were stored in water at 4°C protected from light.

3.4.2 Concentration Determination.

Concentrations of labeled miniature proteins were calculated using the extinction coefficients of the strongest absorbing chromophores. In the case of unlabeled miniature proteins, the extinction was determined directly from the amino acid sequence3. In the case of labeled peptides, fluorescein absorption at 500 nm in 8 M urea, 100 mM Tris-HCl, pH 9.0 was measured using the molar extinction coefficient 78,000 $M^{-1}cm^{-1}$ at 500 nm (Invitrogen). The concentration of rhodamine labeled variants was determined in water using an extinction coefficient of 87,000 $M^{-1}cm^{-1}$. The concentration of SDEX labeled variants were determined using an extinction of 12,000 $M^{-1}cm^{-1}$ at 242 nm in DMSO[171].

3.4.3 Circular Dichroism.

CD spectra of unlabeled peptides were recorded from samples containing 25 μM peptide in PBS as described[62].

3.4.4 Activation of GR by dexamethasone labeled peptides and miniature proteins.

HeLa cells transfected with pK7-GR-GFP for 24 - 36 hr were plated (104/well in 200 µL) the day prior to experiments in 96 well glass bottom plates (Matrical). The media was replaced with HKR buffer containing 300 nM Hoescht and the cells incubated for 30 min at 37°C. The media was then replaced with 150 µL HKR buffer or HKR buffer containing 1 μM dexamethasone labeled cationic miniature protein or peptides or 1 μM dexamethasone as a positive control. The cells were incubated for a further 30 min at

37°C before epifluorescence imaging using a Zeiss Axiovert 200M microscope fitted with a mRM digital camera, a 63 x 1.3NA PlanApo objective and an EXFO-Excite illumination source. Dual color epifluorescence images were acquired using the following filter sets: Hoescht, Zeiss Filter Set #49 (ex G 365 nm, FT 395, em BP 445/50 nm); EGFP, Zeiss Filter Set #44 (ex BP 475/50 nm, FT 500, em BP 530/50 nm).

3.4.5 Image analysis using CellProfiler.

The image analysis algorithm used by CellProfiler is customizable and contained within a 'PIPELINE' that describes a series of image manipulation and evaluation steps, such as reading the image, adjusting the intensity, and identifying objects. Each step has several options. We found that the translocation ratio was generally robust, provided nuclei were accurately identified. Using the image of Hoescht fluorescence and Otsu's method of automatic thresholding, nuclei were identified and segmented using 3 class thresholding. The middle class was assigned to background. Closely approximated objects were separated using 'Laplacian of Gaussian' modeling, and touching nuclei were divided using the 'shape' option. Nuclear regions were expanded by 2 μm to generate an enlarged region. The nuclear region was subtracted from enlarged region to generate a ring surrounding the nucleus (surround). The diameter of the surrounding region was chosen to be 2 μm larger than the nucleus, within the range (1 – 3 μm) used by previously reported high content screens[172, 173]. The ratio of the median intensities of the nuclear and surrounding region of the GFP or mCherry (as appropriate) image were collected for every cell examined. Those cells falling below a threshold of intensity (10% of the maximum intensity of the image) were discarded as 'untransfected'. Typically, 20-60 transfected cells from 15-30 images were examined for each experimental condition.

3.4.6 Examination of Membrane Integrity via Cell Impermeable Dyes.

The loss of membrane integrity was examined via assaying nuclear fluorescence after exposure to Sytox Blue (Invitrogen), a cell impermeant dye that fluoresces brightly when complexed with DNA. Syto Orange 85 (also Invitrogen) was used as a cell permeable counter stain. HeLa cells plated in glass bottom microtitre plates were exposed for 30 min at 37°C to 1 μM dexamethasone labeled cationic miniature protein or peptide in HKR buffer containing 1 μM Sytox Blue and 1 μM Syto Orange. Buffer without peptide or containing 10 μg/mL Saponin was used as negative and positive controls respectively. After incubation cells were washed briefly in HKR buffer, and examined immediately via epifluorescence microscopy (Zeiss Axiovert 200M). Sytox Blue fluorescence was visualized using Filter Set 49, Syto Orange 85 was visualized using Filter Set 43. Images were analyzed using ImageJ.

3.4.7 Competition binding of dexamethasone labeled miniature proteins or peptides with Fluormone® for the glucocorticoid receptor.

Differences in the affinity of dexamethasone-labeled molecules for the human glucocorticoid receptor were assessed via competition binding. For this purpose, we employed a Glucocorticoid Receptor Competitor Assay Kit (Invitrogen). Varying concentrations of dexamethasone-labeled miniature protein or peptide (or dexamethasone as a control) were suspended in 48 μL Screening Buffer (10 mM $NaPO_4$, pH 7.4; 20 mM Na_2MoO_4, 1 mM EDTA, 2% DMSO) and incubated with 2.5 nM FLUORMONE GS1 and 4 nM recombinant human GR for 3 h in the dark at room temperature. Following incubation, the fluorescence polarization of each solution was measured on an Analyst AD plate reader (LJL Biosystems). Each data point represents the average fluorescence polarization ($\pm$ standard deviation) of three separate experiments. Data were plotted as polarization (mP) vs. Log[inhibitor] and fit to the following equation using GraphPad software (Prism):

$$mP = Pmin + \frac{Pmax\text{-}Pmin}{1 + 10^{([C]-Log(IC_{50}))}}$$

where mP is the polarization of the sample, [C] is the concentration of the test compound, Pmax is the polarization of the sample when no competitor is added, and Pmin is the polarization that represents maximum competition.

3.4.8 Peptide Degradation.

Approximately 200,000 HeLa cells were treated with rhodamine labeled cationic miniature proteins and peptides as described for experiments using flow cytometry to measure peptide uptake. These cells were pelleted in 96 well microtitre plates (300*g*, 3 min) and lysed in 40 μL RIPA buffer (150 mM NaCl, 50 mM Tris pH 7.4, 50 mM EDTA pH 7.4, 0.1% SDS, 1% Triton X-100) on ice for 30 min. The insoluble material was cleared from the cell lysate via centrifugation (10,000g, 10 min) and 20 μL of the supernatant was diluted into 400 μL of 15% CH_3CN in water. These samples were analyzed via reverse phase HPLC (Shimadzu Instruments) fitted with a fluorescence detector and a C18 column (Poroshell 120 SB-C18, 2.7µm, 100 mm x 3 mm ID, Agilent; Fig. 3.8) or a C8 column (Kinetex C8, 2.6µm, 100 mm x 2.6 mm ID, Phenomenex; Fig. 3.9).

3.4.9 Resistance to Cathepsin Acid Proteases

50 pmol **aPP**^{R}, **5.2**^{R}, **5.3**^{R}, **Tat**^{R}, or **Arg8**^{R} were incubated with 5 pg cathepsin B, D, or L for 30 min at 37 °C in 10 µL of 20 mM sodium citrate-phosphate buffer (pH 4.7) containing 150 mM NaCl. At the end of the incubation, the reaction was stopped by adding 40 µL of a solution containing a 84:15:1 ratio of water:EtOH:TFA before analysis by HPLC (Shimadzu Instruments) fitted with a fluorescence detector and a C8 column (Kinetex C8, 2.6µm, 100 mm x 2.6 mm ID, Phenomenex).

3.4.10 Generation of Cytoplasmic Extracts using SLO

To generate cytosolic extracts, we first we verified that we could selectively release the cytoplasmic contents of HeLa cells using the published procedures. 2×10^5 HeLa cells were treated with 20 μg/mL SLO in PBS containing 5 mM DTT (20 min, 4 °C), washed (three PBS washes, 150 μL each), and resuspended in 40μL PBS. To allow oligomerization of membrane bound SLO and pore formation, cells were incubated at 37 °C for 15 min. Cells were separated from cytoplasmic extracts by centrifugation (2 min $\times 300g$) and the supernatant (cytoplasmic extract) retained.

To monitor the fraction of cells permeabilized 150μL of PBS containing 1μM propidium iodide (PI), a marker of permeabilized cell membranes, was added to the cell pellet following slow treatment and the cells examined by flow cytometry. Under these conditions, 85 to 89% of the cells accumulated PI, indicating that the plasma membrane had been disrupted. To examine the contents of the cytoplasmic extracts, the insoluble material was cleared from the cytoplasmic extracts via centrifugation (10,000g, 10 min) and 20 μL of the supernatant was diluted into 40 μL of 15% CH_3CN and 0.1% TFA in water. These samples were analyzed via reverse phase HPLC (Shimadzu Instruments) fitted with a fluorescence detector and a C8 column (Kinetex C8, 2.6μm, 100 mm x 2.6 mm ID, Phenomenex). To confirm that these conditions did not release material present within lysosomes, 2×10^5 HeLa cells were treated with SLO at 4 °C, washed, resuspended in PBS, and incubated at 37 °C for 15 min as described above. The cells were then pelleted by centrifugation (2 min $\times 300g$) and the supernatant recovered, after which the cell pellet was treated with detergent lysis buffer (RIPA buffer) to solubilize the membranes. Total cell lysate from an equal number of HeLa cells that were not treated with SLO served as a positive control. These three fractions (total cell lysate, supernatant from SLO treated cells, and the solubilized cell pellet) were then incubated with a solution containing 7.5 mM 4-nitrophenyl-N-acetyl-β-D-glucosaminide (Sigma-Aldrich) in 0.1 M citrate-phosphate buffer, pH 4.7. 4-nitrophenyl-N-acetyl-β-D-glucosaminide is a

chromogenic substrate for the lysosomal enzyme β-hexosamididase which, via hydrolysis, releases 4-nitrophenol whose presence can be spectrophotometrically monitored via absorbance of 405 nm light.

3.4.11 Analytical Ultracentrifugation

Miniature proteins were dissolved in either PBS or HKR buffer at 60, 20, or 10 μM concentration and loaded into 6 chambered cells. Buffer blanks without peptide were used as controls. The cells were loaded into the rotor, and spun in a Beckman XL-1 analytical ultracentrifuge with absorbtion optics at increasing speeds (36,000, 42,000, 50,000, and 60,000 rpm). Peptide sedimentation profiles were assessed at both 280 nm and 230 nm, using a 0.1 mm step size for the scan position. In this way, scans were completed each hour. The approach to equilibrium was monitored using the MATCH function of HETEROANALYSIS, in which the root-mean-square-deviation (rmsd) between successive scans is calculated. Centrifugation at each speed was continued for a minimum of 6 hr, until the rmsd was no longer decreasing, or until the rmsd was < 0.005 between successive scans. At this point to rotational speed was increased or, after reaching equilibrium at 60,000 rpm, the run was stopped.

Data analysis began with the calculation of physical parameters. Solvent density, ρ, was calculated using SEDNTERP, while partial specific volume ($\overline{v}$) was calculated by hand using the method of [163], making adjustments for the inclusion of the c-terminal amide, as well as the presence (or absence) of a dexamethasone coupled to an N-terminal lysine. The extinction coefficient for unlabeled miniature proteins was calculated directly from the amino acid sequence[174]. For dexamethasone labeled molecules, the absorbtion of dexamethasone at 280 nm (863 $M^{-1}\ cm^{-1}$) was added to the value calculated from the amino acid sequence. The absorbance of dexamethasone labeled molecules and unlabeled molecules at 230 nm was measured relative to the absorbance at 280 nm. The molar extinction coefficient at 230 nm for dexamethasone labeled miniature proteins was

determined using the relative absorbances: $\varepsilon_{230} = A_{230} \times \frac{\varepsilon_{280}}{A_{280}}$.

Next scans of sedimentation profiles obtained after the sample arrived at equilibrium were loaded into HETEROANALYSIS. Scan points representing the solution column were selected manually. In some cases, individual points were excluded because of spurious increases in absorbtion. Refitting the data with these points included, led to no change in the calculated association constant of the final model. To examine the presence of a self-association, data were first fit (using non-linear least-squares regression) to a model representing the sedimentation of a single ideal species:

$$C(r) = C(r_0)e^{M(1-\bar{v}\rho)\frac{\omega^2(r^2-r_0^2)}{RT}} + B_0$$

where M is the molecular weight, RT is the product of the universal gas constant and the temperature in kelvin, $C(r_0)$ is the apparent concentration of the sedimenting species at an arbitrary reference point in the solution column (the miniscus), and $C(r)$ is the apparent concentration at radius r, and B_0 is an arbitrary adjustment for the background absorbtion of the solution. The usuable range of absorbance represents the limits of accurate measures of $C(r)$, usually taken to be between 0.1 and 1, acceptable fits generally show rmsd ≤ 0.01 and small B_0, usually < 0.01. A best-fit molecular weight significantly higher than the calculated molecular weight is evidence of self-association.

Next, to examine possible association states of the studied molecule, the data are fit to a general model of reversible two-state (monomer, $n-$mer) self-association, using the observation that at equilibrium, the law of mass action specifies that the concentration of multimer at all points along the column is satisfies $K_a \cdot C_m^n(r) = C_{nM}(r)$, where $C_m(r)$ and $C_{nM}(r)$ are the concentrations of the monomer and $n-$mer, at radius r respectively. The model is specified by the following equation

$$C_{total}(r) = C_m(r_0)e^{M(1-\bar{v}\rho)\frac{\omega^2(r^2-r_0^2)}{RT}} + K_a\left(C_m(r_0)\right)^n e^{nM(1-\bar{v}\rho)\frac{\omega^2(r^2-r_0^2)}{RT}} + B_0$$

in which all variables are specified, with the exception of K_a, B_0 and n.

Finally, the sedimentation profiles are fit to an integer association state using the general two-state monomer-to-n–mer model above, except n is fixed at the integer closest to the best fit n found when n was allowed to vary continuously.

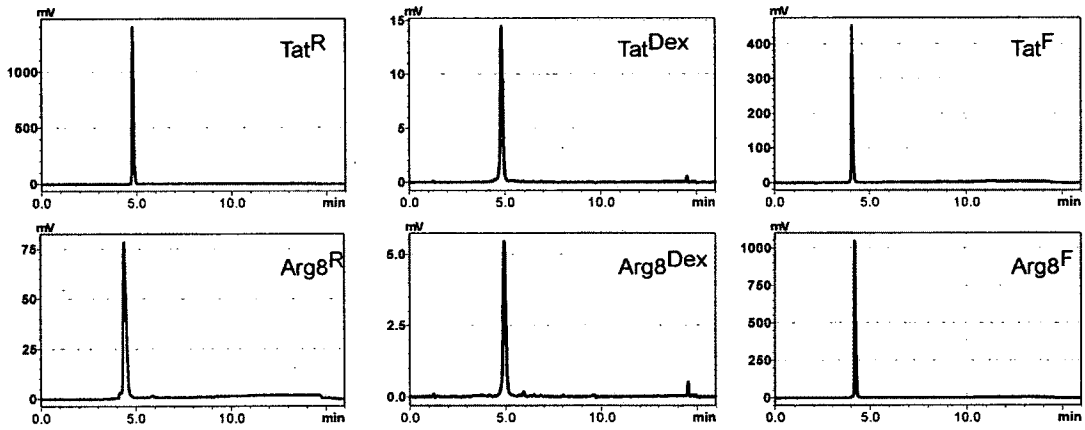


Figure 3.19: HPLC analysis of labeled peptides used in this study.

Name	Sequence	Formula	Calc (+m/z)	Meas
aPP	GPSQPTYPGDDAPVEDLI**R**FYNDLQQYLNVVTRHRYC^{acetyl}-NH2	$C_{195}H_{289}N_{54}O_{61}S_1$	4399	4400
2.1	GPSQPTYPGDDAPVEDLI**R**FYNDLQ**R**YLNVVTRHRYC^{acetyl}-NH2	$C_{196}H_{293}N_{56}O_{60}S_1$	4427	4430
2.2	GPSQPTYPGDDAPVEDLI**R**FY**R**DLQQYLNVVTRHRYC^{acetyl}-NH2	$C_{197}H_{295}N_{57}O_{60}S_1$	4441	4445
3.2	GPSQPTYPGDDAPVEDLI**R**FY**R**DLQ**R**YLNVVTRHRYC^{acetyl}-NH2	$C_{198}H_{299}N_{58}O_{59}S_1$	4469	4475
4.2	GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YLNVVTRHRYC^{acetyl}-NH2	$C_{199}H_{304}N_{61}O_{57}S_1$	4496	4502
4.3	GPSQPTYPGDDAPVEDLI**R**FY**R**DL**RR**YLNVVTRHRYC^{acetyl}-NH2	$C_{199}H_{303}N_{60}O_{58}S_1$	4497	4504
5.2	GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YL**R**VVTRHRYC^{acetyl}-NH2	$C_{201}H_{311}N_{63}O_{56}S_1$	4538	4545
5.3	GPSQPTYPGDDAPV**R**DLI**R**FY**R**DL**RR**YLNVVTRHRYC^{acetyl}-NH2	$C_{200}H_{308}N_{63}O_{56}S_1$	4524	4531
6.3	GPSQPTYPGDDAPV**R**DL**RR**FY**R**DL**RR**YLNVVTRHRYC^{acetyl}-NH2	$C_{200}H_{309}N_{67}O_{57}S_1$	4567	4573
Name	**Fluorescein Labeled Sequences**	**Formula**	**Calc (+m/z)**	**Meas**
aPP^F	GPSQPTYPGDDAPVEDLI**R**FYNDLQQYLNVVTRHRYC^{FLU}-NH2	$C_{215}H_{302}N_{55}O_{65}S_1$	4729	4735
2.1^F	GPSQPTYPGDDAPVEDLI**R**FYNDLQ**R**YLNVVTRHRYC^{FLU}-NH2	$C_{216}H_{306}N_{57}O_{64}S_1$	4758	4763
2.2^F	GPSQPTYPGDDAPVEDLI**R**FY**R**DLQQYLNVVTRHRYC^{FLU}-NH2	$C_{217}H_{308}N_{58}O_{64}S_1$	4777	4778
3.2^F	GPSQPTYPGDDAPVEDLI**R**FY**R**DLQ**R**YLNVVTRHRYC^{FLU}-NH2	$C_{218}H_{312}N_{59}O_{63}S_1$	4800	4806
4.2^F	GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YLNVVTRHRYC^{FLU}-NH2	$C_{219}H_{317}N_{63}O_{61}S_1$	4827	4834
4.3^F	GPSQPTYPGDDAPVEDLI**R**FY**R**DL**RR**YLNVVTRHRYC^{FLU}-NH2	$C_{219}H_{316}N_{62}O_{62}S_1$	4828	4828
5.2^F	GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YL**R**VVTRHRYC^{FLU}-NH2	$C_{221}H_{323}N_{64}O_{60}S_1$	4870	4871
5.3^F	GPSQPTYPGDDAPV**R**DLI**R**FY**R**DL**RR**YLNVVTRHRYC^{FLU}-NH2	$C_{220}H_{321}N_{64}O_{60}S_1$	4855	4855
6.3^F	GPSQPTYPGDDAPV**R**DL**RR**FY**R**DL**RR**YLNVVTRHRYC^{FLU}-NH2	$C_{220}H_{322}N_{68}O_{60}S_1$	4989	4903
Name	**Rhodamine Labeled Sequences**	**Formula**	**Calc (+m/z)**	**Meas**
aPP^R	K^{Rhod}GPSQPTYPGDDAPVEDLI**R**FYNDLQQYLNVVTRHRY-NH2	$C_{223}H_{323}N_{57}O_{66}S_2$	4907	4906
4.2^R	K^{Rhod}GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YLNVVTRHRY-NH2	$C_{227}H_{338}N_{64}O_{61}S_2$	5008	5007
4.3^R	K^{Rhod}GPSQPTYPGDDAPVEDLI**R**FY**R**DL**RR**YLNVVTRHRY-NH2	$C_{227}H_{337}N_{63}O_{62}S_2$	5005	5004
5.2^R	K^{Rhod}GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YL**R**VVTRHRY-NH2	$C_{229}H_{344}N_{66}O_{60}S_2$	5046	5045
5.3^R	K^{Rhod}GPSQPTYPGDDAPV**R**DLI**R**FY**R**DL**RR**YLNVVTRHRY-NH2	$C_{228}H_{342}N_{66}O_{60}S_2$	5032	5031
Tat^R	K^{Rhod}GRKKRRQRRRPPQY-NH2	$C_{112}H_{181}N_{41}O_{24}S_2$	2550	2549
Arg$_8^R$	K^{Rhod}RRRRRRRR-NH2	$C_{81}H_{138}N_{37}O_{15}S_2$	1934	1935
Name	**Dexamethasone Labeled Sequences**	**Formula**	**Calc (+m/z)**	**Meas**
aPP^{Dex}	K^{Dex}GPSQPTYPGDDAPVEDLI**R**FYNDLQQYLNVVTRHRY-NH2	$C_{231}H_{326}F_1N_{55}O_{64}S_1$	4828	4826
4.2^{Dex}	K^{Dex}GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YLNVVTRHRY-NH2	$C_{225}H_{341}F_1N_{62}O_{60}S_1$	4926	4925
4.3^{Dex}	K^{Dex}GPSQPTYPGDDAPVEDLI**R**FY**R**DL**RR**YLNVVTRHRY-NH2	$C_{225}H_{340}F_1N_{61}O_{61}S_1$	4927	4926
5.2^{Dex}	K^{Dex}GPSQPTYPGDDAPV**R**DLI**R**FY**R**DLQ**R**YL**R**VVTRHRY-NH2	$C_{227}H_{347}F_1N_{64}O_{59}S_1$	4968	4968
5.3^{Dex}	K^{Dex}GPSQPTYPGDDAPV**R**DLI**R**FY**R**DL**RR**YLNVVTRHRY-NH2	$C_{226}H_{345}F_1N_{64}O_{59}S_1$	4954	4951
Tat^{Dex}	K^{Dex}GRKKRRQRRRPPQY-NH2	$C_{110}H_{182}F_1N_{39}O_{23}S_1$	2472	2472
Arg$_8^{Dex}$	K^{Dex}RRRRRRRR-NH2	$C_{79}H_{141}F_1N_{35}O_{14}S_1$	1856	1857

Figure 3.20: Mass spectral data obtained for miniature proteins and peptides in this work.

Chapter 4

Cationic Miniature Proteins Reach The Cytoplasm By Escaping From Early Endosomes

4.1 Background

While there is little controversy that under certain conditions some 'cell-penetrating' peptides can reach the cytoplasm of cells when added to the surrounding media, very little is considered accepted regarding how they do so. Some models for cytoplasmic access focus on direct plasma membrane permeation[175], perhaps aided by ion-pair guided diffusion[121, 120], the creation of inverse micelles[106], membrane disruption[176, 125, 142], or peptide self-association[177] with the possible formation of pores[178, 179]. Others advocate that peptides cannot cross the plasma membrane directly, but rather enter the cell through endocytic mechanisms[180]. Within this latter category, there has been much debate on the particular mechanism of endocytosis that is responsible for the internalization of the peptide into endosomes, however, by itself, the mechanism by which peptides arrive in endosomes sheds little light on how escape to the cytoplasm may oc-

cur. Nevertheless, because the field has focused so much on this question, a brief review is pertinent.

4.1.1 Endocytic mechanisms of cationic peptide internalization

The mechanistic understanding of endocytosis has become much more sophisticated in the last 10 years, but with this sophistication comes some confusion. Several recent reviews have restored order through the reinstatement of (in some cases, previously developed[181]) experimental classification schemes[182, 183]. Three types general endocytosis should be considered, the first two being phagocytosis and endocytosis mediated by clathrin or other coat proteins, the third being endocytosis by mechanisms that are not characteristic of the first two types. Whether nascent endosomes require dynamin activity to separate from the plasma membrane is also a useful feature for classification.

Phagocytosis is generally a process limited to phagocytic cells (macrophages, dendritic cells, etc), is stimulated by the presence of large (i.e. $\geq$ 0.5 μM [184]) particles, involves large changes in the distribution of the plasma membrane and is sensitive to cytochalasin-D, an inhibitor of actin polymerization. The process may involve dynamin, but its activity in this case is not due to membrane scission[185], but rather its ability to recruit Rab5 to the phagocytic cup[186, 187]. Because cationic miniature proteins are monomeric in solution, the role of phagocytosis is less important in these studies.

Clathrin-mediated endocytosis is the major mechanism by which coated vesicles mediate endocytosis[188, 189]. The clathrin coat is responsible for the initial deformation of the plasma membrane (a clathrin coated pit), and adaptor proteins recruit cargo into the growing bud. Transferrin is constitutively internalized by this pathway[190, 137] as is activated epidermal growth factor receptor[139]. Fluorescently labeled versions of these proteins have been used as markers for vesicles originating from this pathway. Caveolin is a second protein localized to the interior leaflet of plasma membrane invaginations, and appear to be the preferential site of entry for certain virus, including SV-

40[191, 192, 193], however its role in mediating endocytosis is being reevaluated[194, 195, 196]. The other major coat proteins, COPI, COPII, are generally not active at the plasma membrane[197]. Dynamin is absolutely required for clathrin mediated endocytosis[198, 199].

The study of endocytosis mechanisms capable of internalizing small particles without the participation of clathrin is a growing field, with a rapidly increasing number of example pathways. In general, these pathways are induced by the inhibition of other pathways, making them more difficult to study. Mayor and Pagano[182] suggest segregating these pathways into those that relying on either RhoA, Cdc42, or Arf6, with the latter two being independent of dynamin activity. Importantly all three of these pathways seem to be sensitive to perturbations in the lipid content[200, 201], and in particular depletion of cholesterol[202], from the plasma membrane. Cholesterol depletion decreases (but does not completely inhibit) clathrin mediated endocytosis[203, 204].

Taken together the following experimental results are expected. Small cargo can enter via mechanisms that are (1) dependent on dynamin, (2) dependent on dynamin and cholesterol, or (3) independent of dynamin (and dependent on cholesterol). Clathrin mediated endocytosis falls into the first catagory, and cargo taken up by this route should colocalize with fluorescently labeled transferrin or EGF. The latter two catagories segregate RhoA and Arf6 pathway (2) from the Cdc42 pathway (3). Inhibition of dynamin clathrin mediated endocytosis may stimulate clathrin independent pathways, so this must be kept in mind when interpreting experimental findings. Peptide uptake may occur by one or several of these mechanisms.

4.1.2 Trafficking of endocytosed peptide

While there may be a multitude of internalization pathways, a great simplification is achieved by the finding that the itineraries of these pathways converge. In the case of clathrin, once the clathrin coated pit pinches off from the plasma membrane, the

phosphoinositides on the membrane are rapidly metabolized from phosphoinositol-4,5-bisphosphoate ($PI(4,5)P_2$) to PI3P, and leading to the recruitment of the small GTPase Rab5[205]. Rab5+ endosomes undergo homotypic fusion (i.e. fusion with other Rab5+ vesicles) to form the compartment known as the 'early endosome'[206, 207]. While the molecular events regarding Rab5 aquisition are less well understood in the case of other endocytic pathways, the cargo internalized by the Cdc42 dependent[208] and Arf-6 dependent[209, 210] pathways are delivered to Rab5+ vesicles. This has led to the current hypothesis that all endocytosed cargo is delivered to the Rab5+ early endocytic compartment.

The early endosomal compartment serves as a cargo sorting station[181]. Sorting molecular cargo destined to return to the cell surface takes place by actin dependent spatial segregation into specialized domains within the early endosomes[211]. These domains then bud off to form tubular/vesicular structures that fuse with the plasma membrane, recycling the cargo back to the cell surface. In the absence of material being recycled to the plasma membrane, early endosomes exchange Rab5 for Rab7[212, 213] forming 'late endosomes', which then fuse with lysosomes[140, 214]. Cargo destined for other cellular locales travel through late endosomes, but the mechanism by which they are selectively delivered to their destinations is not understood.

4.1.3 Reaching the cytoplasm

As stated above cationic miniature proteins may reach the cytoplasm by crossing the plasma membrane directly or by escape from an intracellular vesicle. The ability of cationic miniature protein 5.3 to access the cytoplasm in the absence of endocytosis would strongly suggest the ability to cross the plasma membrane directly. In the absence of this ability, the question arises, from which endosomal population does 5.3 escape? It would also be useful to compare the identify to the answer for 5.3 to the vesicles from which other peptides can escape to the cytoplasm.

The molecular character of endosomes changes with their maturation. As endosomes mature, they increase in byouant densitiy, acquire different lipid and protein components, and increase their lumenal concentration of acid. The ability to reach the cytoplasm may be induced by one or more of these bulk physicochemical properties, or it may be due to the recruitment of a protein ligand that facilitates the specific exist of 5.3 or another cell penetrating peptide.

4.2 Results

Three limiting models have been invoked to explain the trafficking of cationic peptides and proteins across the plasma membrane and into the cytosol. The first posits that high local concentrations of positively charged peptides cause membrane disruption[125]. We previously confirmed that the integrity of the cell membrane is not disrupted in the presence of 1 μM cationic miniature protein (see Fig. 3.4, on page 46). A second model invokes ion-pair guided passive membrane diffusion[121, 120]; while the third model invokes endocytosis followed by endosomal release (*vide infra*). We sought to distinguish between remaining models in two ways. First, we asked whether rhodamine labeled cationic miniature proteins are taken up into the same vesicles as transferrin, a classic marker of endocytosis; finally, we asked whether inhibitors of endocytosis block the uptake of these same molecules.

4.2.1 Cationic miniature proteins traffic first into endocytic vesicles

To distinguish between the two remaining mechanisms of uptake, we began by treating HeLa cells with rhodamine labeled miniature proteins in the presence of AlexaFluor-488-transferrin (Tf^{488}) and quantifying colocalized fluorescence using confocal microscopy (Fig. 4.1). Transferrin, a protein whose trafficking serves as a predominant model for the study of endocytosis[183] is rapidly internalized from the plasma membrane into endocytic vesicles[139, 215]. Thus, the localization of rhodamine labeled miniature pro-

tein in transferrin positive vesicles would suggest that those vesicles were endocytic in nature, originating from the plasma membrane. We incubated HeLa cells at 37°C for 30 min with Tf^{488} (25 nM) and 1 μM $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, $\mathbf{Tat}^R$ or $\mathbf{Arg8}^R$. The cells were washed with media, and labeled with Hoescht (to visualize DNA). Images were then obtained by spinning disk confocal microscopy without fixing the cells in order to avoid the possible release of peptide from intracellular vesicles that can occur following fixation by treatment with formaldehyde or methanol[127, 140]. As expected, when excited with 488 nm light, HeLa cells treated with Tf^{488} showed small, discrete areas of green fluorescent signal distributed throughout the cytosol. Addition of $\mathbf{aPP}^R$ led to little or no red fluorescent signal and little or no change in the green signal due to Tf^{488}, confirming earlier results (Fig. 3.3) that $\mathbf{aPP}^R$ is not taken up efficiently by HeLa cells. In contrast, HeLa cells treated with 1 μM $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, $\mathbf{Tat}^R$ or $\mathbf{Arg8}^R$ showed bright red fluorescent puncta throughout the cytosol in a distribution that overlapped with the green fluorescence seen with Tf^{488} and at levels that qualitatively reproduce the trends detected by flow cytometry (Fig. 3.3). With one exception, the fluorescent signal from Tf^{488} and the rhodamine labeled molecules were highly correlated ($R_{488,Rhodamine}$ = 0.619 - 0.779), suggesting that transferrin and cationic miniature proteins/peptides $\mathbf{4.2}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, $\mathbf{Tat}^R$ or $\mathbf{Arg8}^R$ are taken up into the same endocytic compartments. Miniature protein $\mathbf{4.3}^R$ is also taken up into a transferrin+ compartment, but the correlation ($R_{488,Rhodamine}$ = 0.493) is lower, possibly because the uptake is low (Fig.3.3).

To assess whether peptide trafficking is affected by fluorophore identity, we directly compared the intracellular distributions of rhodamine and fluorescein tagged versions of **aPP**, **4.2**, **4.3**, **5.2**, and **5.3**, as well as **Tat** and **Arg8** (Fig. 4.2). While HeLa cells treated with 1 μM $\mathbf{aPP}^F$ and $\mathbf{aPP}^R$ exhibited little or no fluorescence, those treated simultaneously with rhodamine and fluorescein labeled variants of either **4.2**, **4.3**, **5.2**, or **5.3** contained bright puncta with substantial correlation between the rhodamine and fluorescein signal. Overlap was also observed after treatment with rhodamine and fluorescein labeled versions of **Tat** as well as **Arg8**. The per-pixel correlation between rhodamine and fluo-

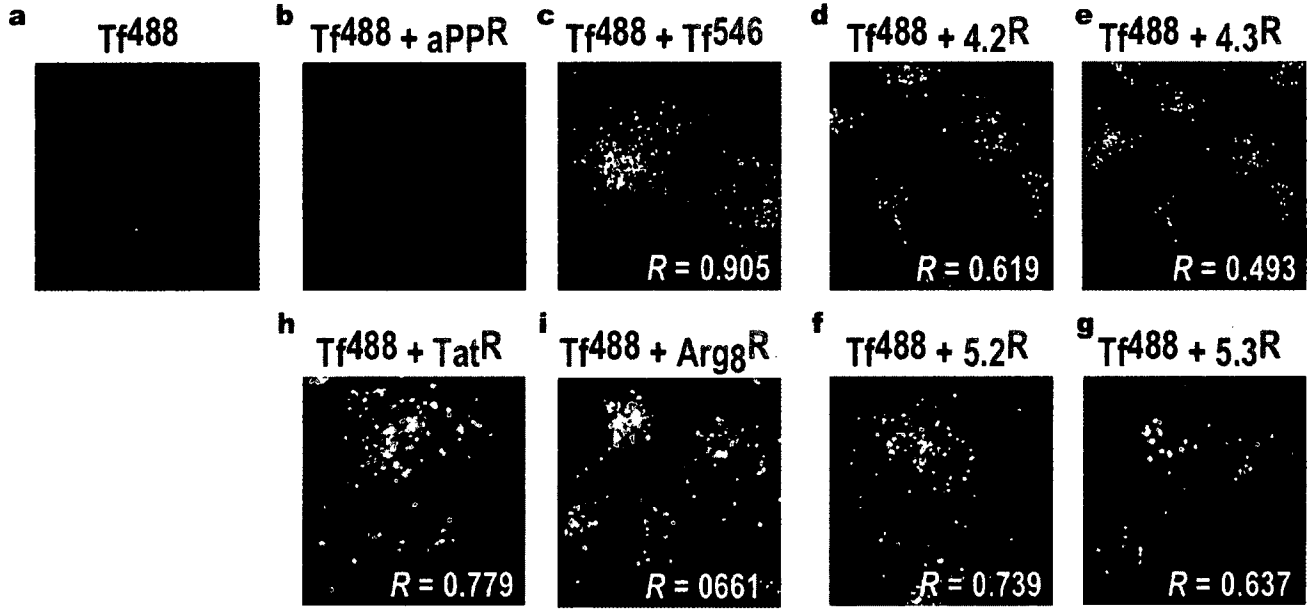


Figure 4.1: Colocalization of alexaFluor-488-transferrin Tf^{488} with rhodamine labeled miniature proteins. Perfect colocalization is characterized by a Pearson's R value (R) equal to 1, while R values near 0 represent little or no colocalization. The correlation value observed when cells were treated with both Tf^{488} and alexa-fluor-546-transferrin is 0.905. Rhodamine labeled cationic miniature proteins and alexa-546-transferrin are shown in red, Tf^{488} is shown in green, Hoescht 33342 (nucleus) is shown in blue.

rescein intensities was found to be highly significant for all cationic miniature proteins ($R > 0.68$) as well as for Arg8 ($R = 0.756$), and Tat ($R = 0.631$). These results confirm that these molecules traffic to similar locations in the cell regardless of fluorescent label location or identity, and recapitulate the similarity in the quantitative analysis of peptide uptake above (Fig. 3.3, on page 43).

Next, we asked whether small molecules known to inhibit endocytosis would block the uptake of rhodamine labeled miniature proteins. Dynasore[216]is an inhibitor of the GTPase dynamin, a protein whose activity is required for clathrin mediated endocytosis (CME)[217], pinocytosis[218, 219], as well as caveolin mediated endocytosis[191]. Depleting cellular cholesterol through treatment with methyl-β-cyclodextrin (MβCD) also inhibits these three processes[191, 204]. Endocytic pathways that continue despite loss of dynamin activity and depletion of cellular cholesterol[217] depend on actin remodeling, a process inhibited by addition of N-ethyl-isopropyl amiloride (EIPA)[220]. Addition of EIPA does not affect the uptake of CME ligands such as transferrin[220]. To test the involvement of these pathways in the uptake of cationic miniature proteins, we pretreated HeLa cells for 30 min with 80 μM dynasore, 5 mM MβCD or 50 μM EIPA before adding 1 μM $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, $\mathbf{Tat}^R$, $\mathbf{Arg8}^R$ or $\mathbf{aPP}^R$. Peptide uptake was allowed to proceed for 30 min at 37°C in the presence of inhibitor; the cells were subsequently washed and visualized by confocal microscopy (Fig. 4.3). The presence of dynasore completely blocked the uptake of $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, as well as $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$; suggesting these molecules are taken up in a dynamin-dependent fashion. MβCD completely blocked the uptake of $\mathbf{4.2}^R$, $\mathbf{4.3}^R$, $\mathbf{5.2}^R$, $\mathbf{5.3}^R$, and $\mathbf{Tat}^R$, but not $\mathbf{Arg8}^R$. Surprisingly cellular uptake of Arg8R was increased in the presence of MβCD, leading to diffuse fluorescence throughout the cytosol, a pattern not seen in the absence of the inhibitor (see Fig. 4.1 and Fig. 4.3). The presence of EIPA also dramatically reduced the uptake of cationic miniature proteins, as well as $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$. The results with $\mathbf{Arg8}^R$ notwithstanding, these data support the contribution of various endocytic pathways to the uptake of cationic miniature proteins and peptides, and not a model based on passive diffusion.

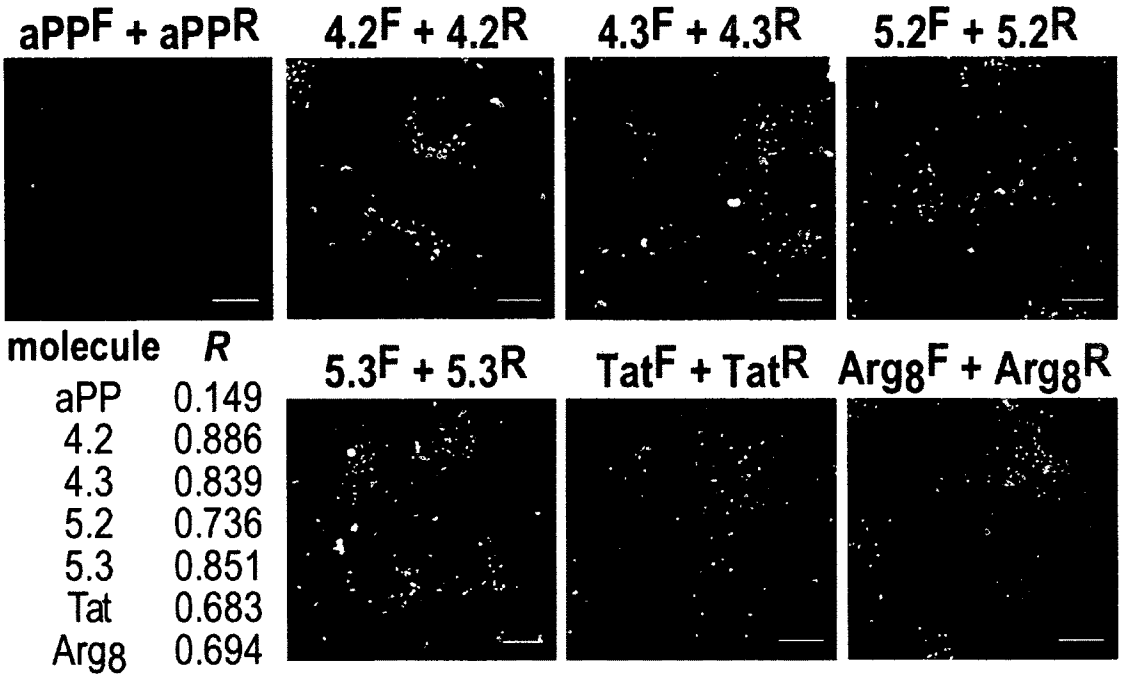


Figure 4.2: Colocalization of rhodamine and fluorescein labeled cationic miniature proteins. HeLa cells were treated with 1 μM rhodamine and fluorescein cationic miniature protein or peptide for 30 min at 37°C before being washed and imaged using confocal microscopy. HeLa cells treated with $\mathbf{aPP}^R$ and $\mathbf{aPP}^F$ show low fluorescence. The remaining molecules show high fluorescence, and molecules sharing the same sequence (but different labels) show high degrees of overlap, with $R > 0.68$. Scale bars are 20 μm.

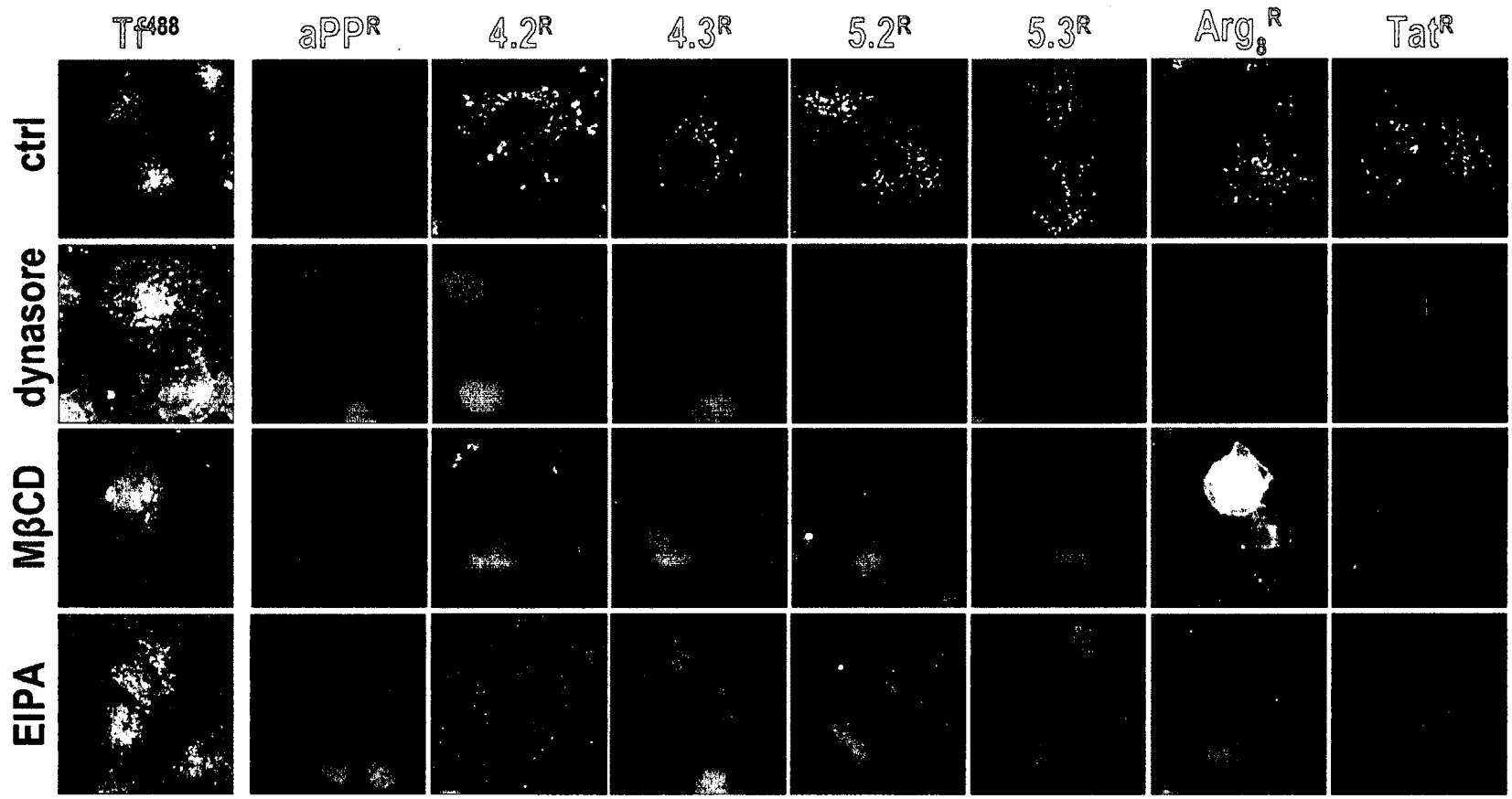


Figure 4.3: Uptake of cationic miniature proteins is inhibited by treatments the block endocytosis. HeLa cells were preincubated in media (ctrl) or media containing 80 μM dynasore, 5 mM MβCD, or 50 μM EIPA before the addition of 1 μM of the indicated rhodamine labeled peptide or Tf^{488} for 30 min. Cells were then washed, labeled with Hoescht (nucleus, blue) and visualized without fixation by confocal microscopy.

4.2.2 Cytoplasmic access requires active endocytosis

In many cases, cationic proteins/polypeptides showing cellular uptake by flow cytometry remain trapped in endosomes and fail to reach the cytosol[128]. This raises the possibility that cytoplasmic access does not depend on endosomal escape, but rather direct diffusion through the plasma membrane. Having shown that cationic miniature proteins enter endocytic vesicles, we sought to test whether their cytoplasmic accumulation was the result of endosomal escape or diffusion acros the plasma membrane. We were mindful of the fact that exchanging the fluorescent rhodamine label for dexamethasone could alter the physical properties of the molecule as well as the manner in which it trafficked in cells. Because inhibitors of endocytosis prevented the entry of cationic miniature proteins into endocytic vesicles, we wondered whether treatment with these same inhibitors would also block cytoplasmic access. Such a finding would support the conclusion that diffusion across the plasma membrane, in the absence of active endocytosis, is an ineffective route to the cytoplasm.

Having identified 5.3^{Dex} as a cationic miniature protein showing significant cytoplasmic access, and hence significant increases the nuclear-to-cytoplasmic translocation ratio of GR-GFP, we set out to determine the pathway by which cytosolic access was achieved. Having previously shown that pharmacologically inhibiting endocytosis blocked uptake of 5.3^{R} (Fig. on the preceding page), we first asked whether inhibiting endocytosis in the same manner also blocked the ability of 5.3^{Dex} to reach the cytoplasm. HeLa cells transfected with GR-GFP were pre-treated for 30 min with the aforementioned inhibitors prior to the addition of either 1 μM dexamethasone (positive control), $\mathbf{aPP}^{Dex}$ (negative control), 5.3^{Dex}, $\mathbf{Tat}^{Dex}$, or $\mathbf{Arg8}^{Dex}$ for 30 min at 37°C (Fig. 6). The cells were then washed and imaged to measure the translocation ratio. None of the endocytosis inhibitors altered the GR-GFP translocation ratio calculated for cells treated with $\mathbf{aPP}^{Dex}$, which remained low after treatment. However, all three inhibitors reduced, to background levels, the GR-GFP translocation ratio calculated after treatment with 5.3^{Dex}, $\mathbf{Tat}^{Dex}$, or $\mathbf{Arg8}^{Dex}$.

Neither EIPA nor dynasore altered the increase in the translocation ratio calculated after treatment with dexamethasone, but MβCD reduced this increase by 53%. This effect may be due to direct complexation of free dexamethasone by MβCD[221].

Because endosomes contain very small volumes, their internal pH is exquisitely sensitive to the activity of proton pumps, including V-ATPases; acidification begins almost immediately upon scission of endosomes from the plasma membrane, as their lumen no longer communicates with the surrounding media. Bafilomycin is a potent inhibitor of the V-type ATPase[222, 223], and its addition to culture media prevents endosomal acidification[123, 222]. To ask whether low vesicular pH was required for cytoplasmic escape, we pretreated HeLa cells expressing GR-GFP with 200 nM bafilomycin for 1 h before exposure to 1 μM dexamethasone, $\mathbf{5.3}^{Dex}$, $\mathbf{Tat}^{Dex}$, $\mathbf{Arg8}^{Dex}$ or $\mathbf{aPP}^{Dex}$ (Fig. on the next paged). Treatment with bafilomycin did not alter the effect of dexamethasone or $\mathbf{aPP}^{Dex}$ on the translocation ratio, consistent with the understanding that dexamethasone freely diffuses into the cytoplasm and that $\mathbf{aPP}^{Dex}$ never enters endocytic vesicles. The increase in translocation ratio seen after cells were exposed to $\mathbf{5.3}^{Dex}$, however, was completely blocked by the presence of bafilomycin ($p = 1.4 \times 10^{-11}$). Bafilomycin also blocked the increase in translocation ratio seen after treatment with $\mathbf{Tat}^{Dex}$ ($p = 5.3 \times 10^{-11}$) or $\mathbf{Arg8}^{Dex}$ ($p = 8.2 \times 10^{-6}$). Thus, while $\mathbf{5.3}^{Dex}$, and to a slightly lesser extent $\mathbf{Tat}^{Dex}$ and $\mathbf{Arg8}^{Dex}$, induce robust translocation of GR-GFP, they fail to do so in the absence of endosome acidification. This finding further supports the model that these molecules do not penetrate the plasma membrane directly, but rather escape to the cytoplasm from acidified endocytic vesicles.

4.2.3 Escape to the Cytoplasm from Early Endosomes

Cell surface protein and lipids are constitutively internalized to form cytoplasmic endosomes. One canonical route, typified by transferrin and its receptor, is internalization via clathrin–coated pits, after which the surrounding vesicle membrane under-

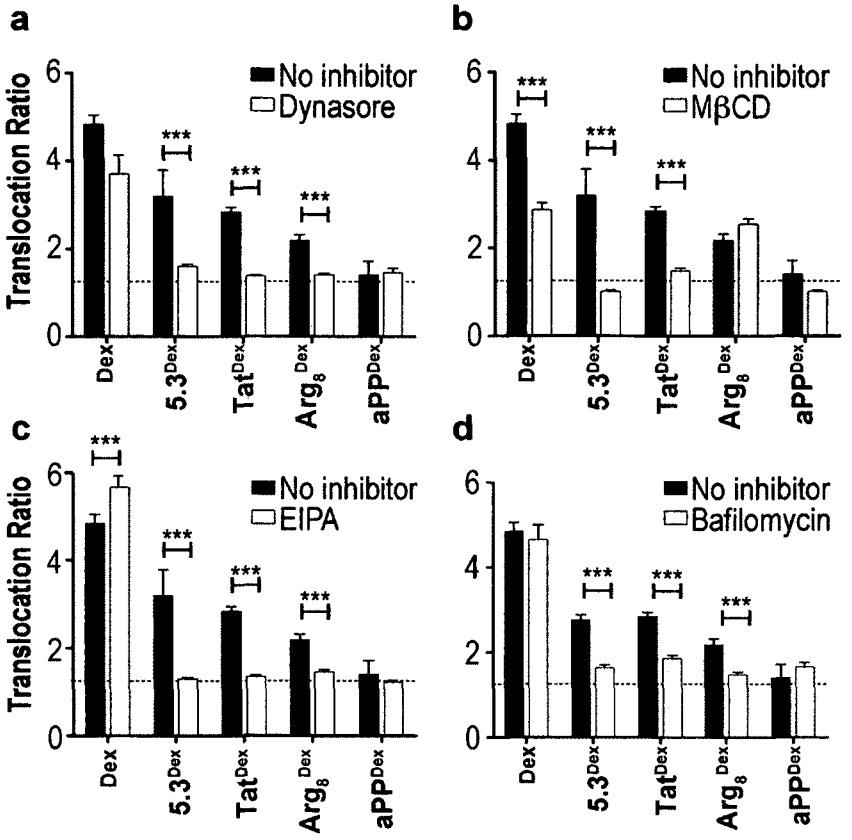


Figure 4.4: Inhibition of endocytosis blocks cytoplasmic access. Translocation of GR-GFP after treatment with 1 μM Dex, 5.3Dex, TatDex, Arg8Dex, or aPPDex in the presence (gray) or absence (black) of various small molecule inhibitors. Inhibitors of endocytosis included (**a**) 80 μM dynasore, (**b**) 5 mM methyl-β-cyclodextran (MBCD), (**c**) 50 μM EIPA. (**d**) Translocation ratio after treatment with cationic miniature proteins and 200 nM bafilomycin . $*p \leq 0.05$; $***p \leq 0.001$; ANOVA with Bonferroni post test.

goes rapid modification resulting in dissociation of the clathrin coat and acquisition of Rab5[215, 205]. The presence of Rab5 recruits cellular factors required for vesicle fusion and maturation, including the phosphotidyl inositol 3-OH kinase, Vps34 [224]. The resulting early endocytic compartment mixes via homotypic fusion with other Rab5+ vesicles[225, 226], and delivers cargo to other cellular locales through the budding off of transport vesicles[211] or Rab conversion[212, 213]. While transferrin is recycled to the cell surface, other cargo, including low density lipoprotein, epidermal growth factor (EGF)[227, 228] and several types of viruses[140, 229, 196] are delivered to late endosomes, marked by Rab7, for eventual degradation in lysosomes[140, 227]. To characterize the intracellular route taken by cationic miniature proteins to the cytoplasm, we first looked for overlap of 5.3^R, $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$ with markers of endocytic uptake and GFP tagged Rab proteins. Then, using small molecule inhibitors and dominant negative Rab variants, we tested the cellular activities required for 5.3^{Dex}, $\mathbf{Tat}^{Dex}$ and $\mathbf{Arg8}^{Dex}$ to enter the cytoplasm.

To characterize the pathway by which 5.3^{Dex} moves from the exterior of the cell to the cytosol, we began with the observation that 5.3^R colocalizes with transferrin (Fig. 4.1), a substrate known to internalize into Rab5+ vesicles[215]. To ask whether 5.3^R was also present in Rab5+ vesicles, HeLa cells were transfected with GFP-Rab5 and treated for 30 min with 1 μM 5.3^R. When these cells were examined by confocal microscopy, 66% of the rhodamine signal overlapped with the signal from GFP-Rab5, confirming that 5.3^R is present in Rab5+ vesicles (Fig. 4.5a). To test whether trafficking of 5.3^R could be arrested at the Rab5 stage, we overexpressed a GTPase-inactive Rab5 mutant, Rab5^{Q79L}[225], that blocks delivery of cargo to late endosomes and arrests vesicle maturation[212, 229, 140, 225]. Observation of HeLa cells overexpressing Rab5^{Q79L}-GFP and treated with 5.3^R showed that nearly all (99%) of the miniature protein localized to enlarged GFP+ endosomes (Fig. 4.5b), suggesting that arresting early endosome maturation also arrests trafficking of 5.3^R at the Rab5+ stage. Similar results were seen with $\mathbf{Arg8}^R$ and $\mathbf{Tat}^R$ (Fig 4.5d-i). Because Rab5 vesicles rapidly deliver their cargo to down-

stream vesicles[212], we also evaluated colocalization of 5.3^R with Rab7-GFP (Fig. 4.5c). HeLa cells transfected with Rab7-GFP and treated with 5.3^R as above, showed that a large fraction (87%) of the rhodamine signal was located in the Rab7-GFP compartment, confirming that 5.3^R enters both early (Rab5+) and late (Rab7+) endosomes. We found no overlap between 5.3^R, $\mathbf{Tat}^R$, or $\mathbf{Arg8}^R$ and galT-GFP[230, 231], a marker of the golgi (see Fig. 4.6).

To ask whether 5.3 uses intracellular trafficking pathways that are distinct from the pathway used by miniature proteins that do not reach the cytoplasm, we directly compared the intracellular trafficking of 5.3 to other cationic miniature proteins, as well as to **Tat** and **Arg8** (Fig. 4.7). Examiniation of HeLa cells treated simultaneously with 1 μM 5.3^R and 4.2^F, 4.3^F, or 5.2^F showed strong correlations ($R \geq 0.68$) between the fluorescence intensity due to the two miniature proteins present in the sample. There was also strong correlation between the fluorescence intensity of 5.3^F and $\mathbf{Arg8}^R$ when applied to the same sample ($R = 0.631$), suggesting that 4.2, 4.3, 5.2, 5.3, and **Arg8**, all occupy the same or similar sets of intracellular compartments. This conclusion was also supported by the observation that >90% of total integrated fluorescence from 5.3 overlapped with the other miniature protein or peptide present in the sample.

The case was different for the comparison of 5.3^R and $\mathbf{Tat}^F$: the correlation between the per-pixel fluorescence intensities was lower ($R = 0.554$). Substantial amounts of green fluorescence from $\mathbf{Tat}^F$ remained visible on the cell surface (arrowheads, Fig. 4.7d), whereas no red fluorescence from 5.3^R was visible at these same locations. In addition, there is a noticeable absence of the small fluorescent vesicles containing 5.3 from the zoomed image that appear in images from the other samples. In contrast, small vesicles are present in the image field, however they contain only $\mathbf{Tat}^F$ and not 5.3^R. Consequently, only 32.1% of the green fluorescence originates from locations occupied by 5.3^R, whereas 92.1% of the red fuorescence originates from locations occupied by $\mathbf{Tat}^F$. Thus, the addition of $\mathbf{Tat}^F$ reduces the presence of 5.3^R is small (but not large) endocytic vesicles. This may be due to competition for a limited number of binding sites on the cell

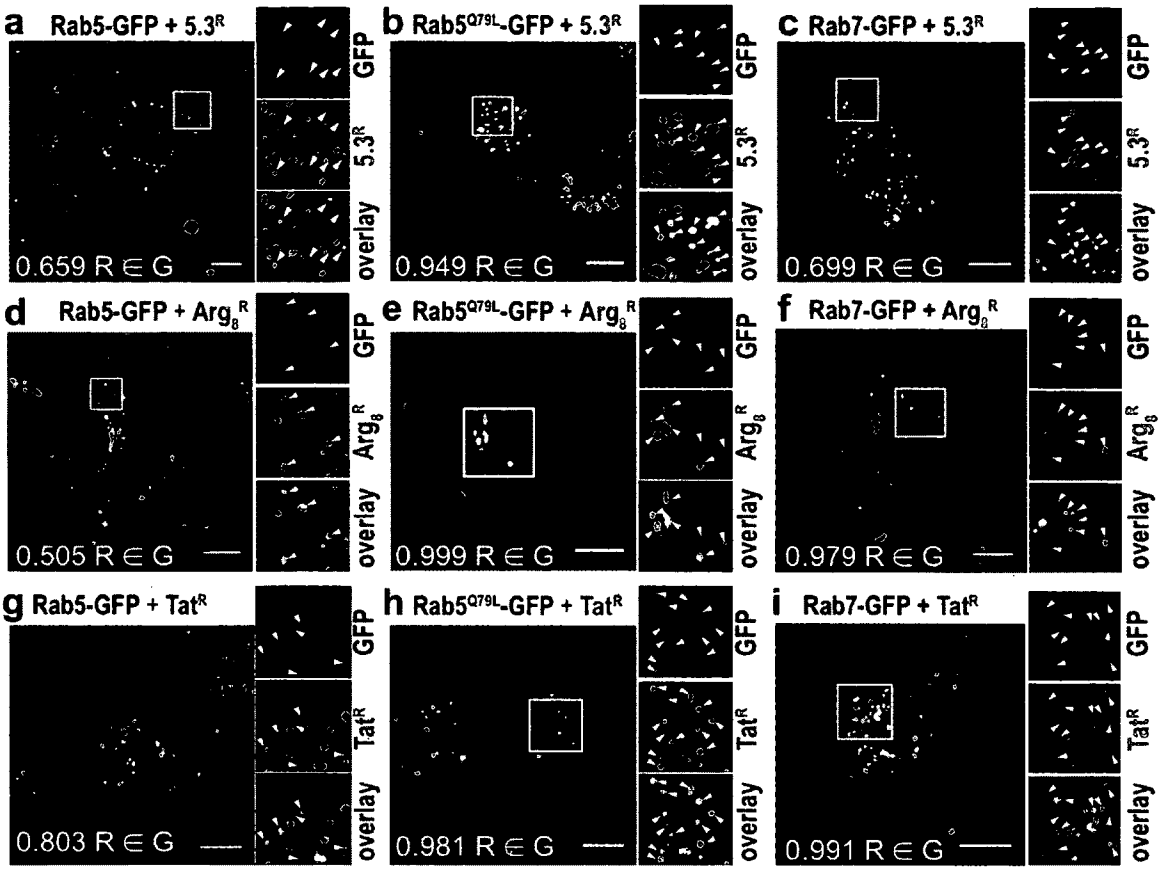


Figure 4.5: Miniature protein 5.3^R enters via endocytosis into Rab5+ vesicles before trafficking to Rab7+ vesicles. HeLa cells transfected with the indicated GFP fusion protein (panels **a** - **i**) were treated with 1 μM 5.3^R (panels **a**-**c**), $\mathbf{Tat}^R$ (panels **d**-**f**) or $\mathbf{Arg8}^R$ (panels **g**-**i**) before being washed and imaged by confocal microscopy. Colocalization of 5.3^R, $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$ with Rab5-GFP is moderate (panels **a**, **d**, **g**) but can be increased by arresting Rab5 maturation via overexpression of $Rab5^{Q79L}$-GFP. (panels **b**, **e**, **h**). 5.3^R, $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$ are delivered to Rab7+ endosomes (panels **c**, **f**, **i**).

surface.

These data suggest that during the first 30 min after treatment, 5.3^R, $\mathbf{Tat}^R$, and $\mathbf{Arg8}^R$ follow a shared path through Rab5+ and then Rab7+ vesicles and provide a starting point to ask which, if any, of these trafficking events are required to reach the cytoplasm.

To identify the point or points along the endocytic pathway at which these three molecules escape to the cytsosol, we blocked vesicle maturation by overexpressing $Rab5^{Q79L}$-GFP and assayed for GR activation following treatment with 5.3^{Dex}, $\mathbf{Tat}^{Dex}$, and $\mathbf{Arg8}^{Dex}$ (Fig. 4.8). In this setting cytosolic localization was assessed *via* dual color microscopy

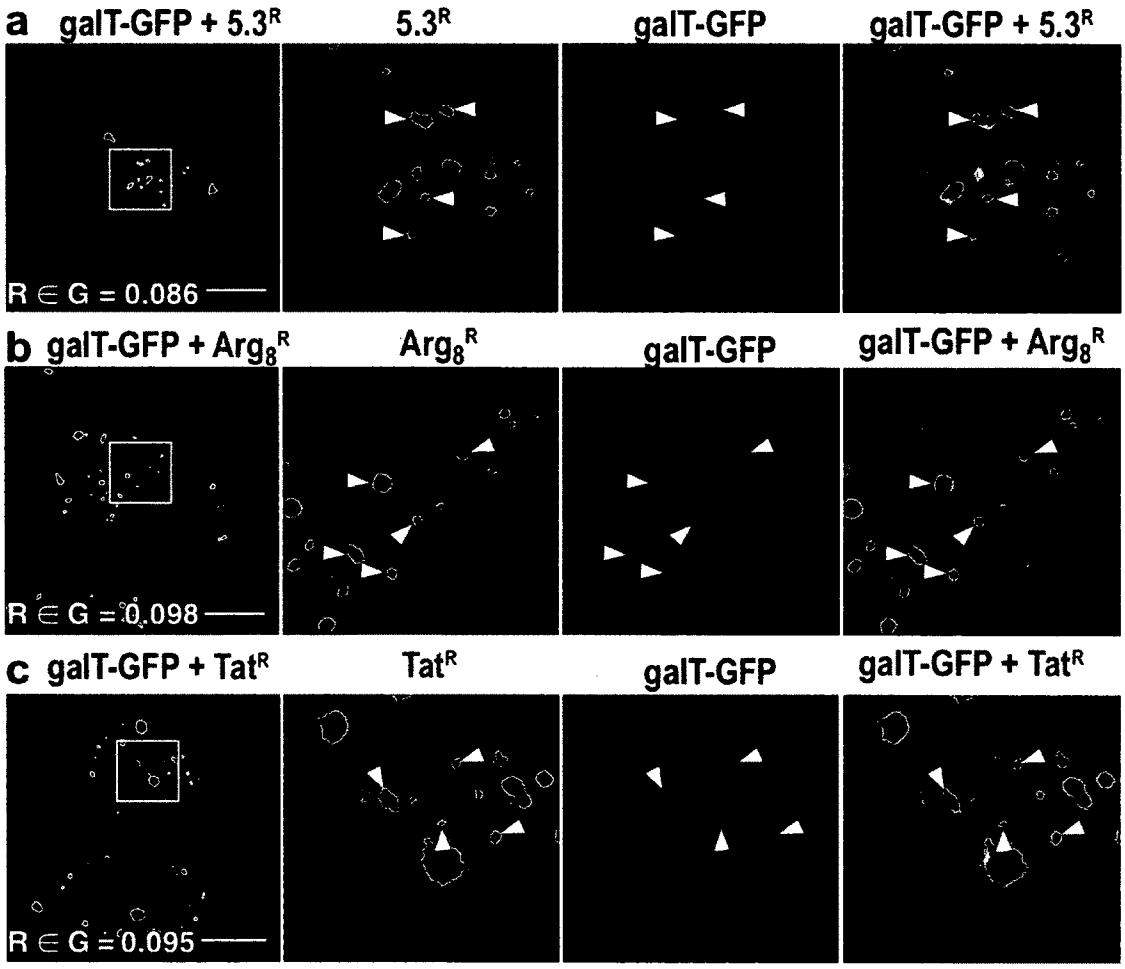


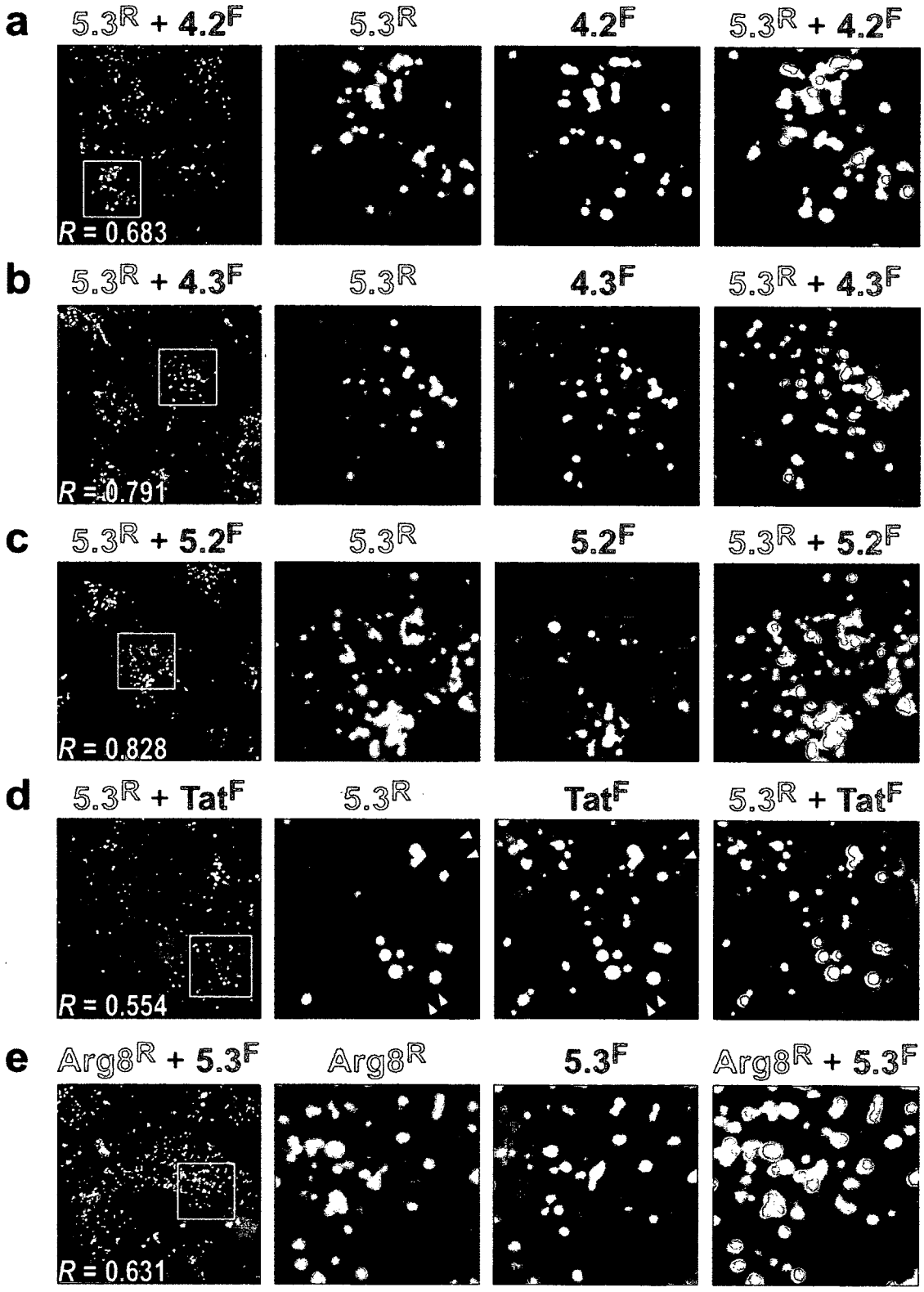
Figure 4.6: Cell permeable cationic miniature proteins are not found within the Golgi. HeLa cells were transfected with galactosyltransferase tagged to GFP, (a marker of the Golgi) before plating in 96 well glass bottom plates and growth overnight. The media was replaced with HKR buffer containing (a) 1 μM 5.3^R, (b) Tat^R or (c) $Arg8^R$ and incubated at 37°C for 30 min, after which the cells were washed and imaged. No red fluorescence (arrowheads) was found within Golgi areas identified by high levels of green fluorescence. The fraction of rhodamine signal overlapping green signal ($R \in G$) within the golgi region (boxed in white) is low (less than 0.1). Selected vesicles showing red fluorescence are indicated with white arrowheads. None overlap with green fluorescence signal. Scale bars are 20 μm.

using an orthogonally labeled GR-mCherry. This experiment allowed us to ask whether inhibiting Rab5 vesicle maturation also blocked the ability of $\mathbf{5.3}^{Dex}$, $\mathbf{Tat}^{Dex}$, and $\mathbf{Arg8}^{Dex}$ to reach the cytosol. Control experiments reveals that the observed translocation ratio of untreated HeLa cells expressing $Rab5^{Q79L}$-GFP or Rab5-GFP and GR-mCherry is near unity, as expected (Fig. 4.8a). This remained unchanged after treatment 30 min with 1 μM $\mathbf{aPP}^{Dex}$ (translocation ratio = 1.32 ± 0.07, 1.25 ± 0.04, respectively); however, these values increased substantially after treatment for 30 min with 1 μM of dexamethasone (Rab5-GFP expressing cells, 2.24 ± 0.15; $Rab5^{Q79L}$-GFP expressing cells, 2.45 ± 0.21; Fig. 4.8a). Treatment with $\mathbf{5.3}^{Dex}$ results in translocation ratios that are similar regardless of whether cells were transfected with GTPase-inactive $Rab5^{Q79L}$ or wild-type Rab5. Analogous findings are observed when cells were treated with $\mathbf{Arg8}^{Dex}$: the translocation ratios calculated are independent of Rab5 activity (1.59 ± 0.12 *vs* 1.60 ± 0.09). By contrast, overexpression of $Rab5^{Q79L}$ blocked the ability of $\mathbf{Tat}^{Dex}$ to reach the cytoplasm (1.26 ± 0.03 *vs* 1.57 ± 0.04, $p = 0.0057$) indicating that **Tat** escapes inefficiently from Rab5+ vesicles. These findings suggest that **5.3** and **Arg8** escape to the cytoplasm from Rab5+ vesicles, but **Tat** escapes from later along the endocytic pathway.

To confirm the finding that Rab5 vesicle maturation is dispensable for cytoplasmic access, we treated cells with 200 nM wortmannin, a pharmacologic inhibitor of phosphoinositide-3-OH kinase (PI3K) and an agent widely used to block the maturation

---

Figure 4.7 *(following page)*: Cationic miniature protein **5.3** shares an overlapping intracellular distribution with other cationic miniature proteins, **Tat** and **Arg8**. HeLa cells treated with $\mathbf{5.3}^{R}$ and (**a**) $\mathbf{4.2}^{F}$, (**b**) $\mathbf{4.3}^{F}$, (**c**) $\mathbf{5.2}^{F}$, or (**d**) $\mathbf{Tat}^{F}$ labeled cationic miniature proteins show high degrees of overlap. A single confocal slice is shown (*left*). Zoomed images (boxed white area on the wide-field image) are shown for individual fluorescence channels and an overlay (*right*). (**e**) $\mathbf{5.3}^{F}$ also shows substantial overlap with $\mathbf{Arg8}^{R}$. Pearson's correlation coefficient (R) calculated between red and green pixel intensities is shown. The correlation is high ($R \geq 0.63$) for all cases except for $\mathbf{5.3}^{R}$ and $\mathbf{Tat}^{F}$ ($R = 0.554$). In this case, there is an absence of fluorescence from **5.3** in small vesicles that is seen in the other images. $\mathbf{Tat}^{F}$ also shows fluorescence remaining on the plasma membrane (arrowheads), whereas $\mathbf{5.3}^{R}$ does not. Calculation of overlap using Mander's coefficient reveals that 92.1% of the red fluorescence overlaps green, whereas only 32.1% of green fluorescence overlaps red.


a
5.3R + 4.2F
5.3R
4.2F
5.3R + 4.2F
R = 0.683
b
5.3R + 4.3F
5.3R
4.3F
5.3R + 4.3F
R = 0.791
c
5.3R + 5.2F
5.3R
5.2F
5.3R + 5.2F
R = 0.828
d
5.3R + TatF
5.3R
TatF
5.3R + TatF
R = 0.554
e
Arg8R + 5.3F
Arg8R
5.3F
Arg8R + 5.3F
R = 0.631

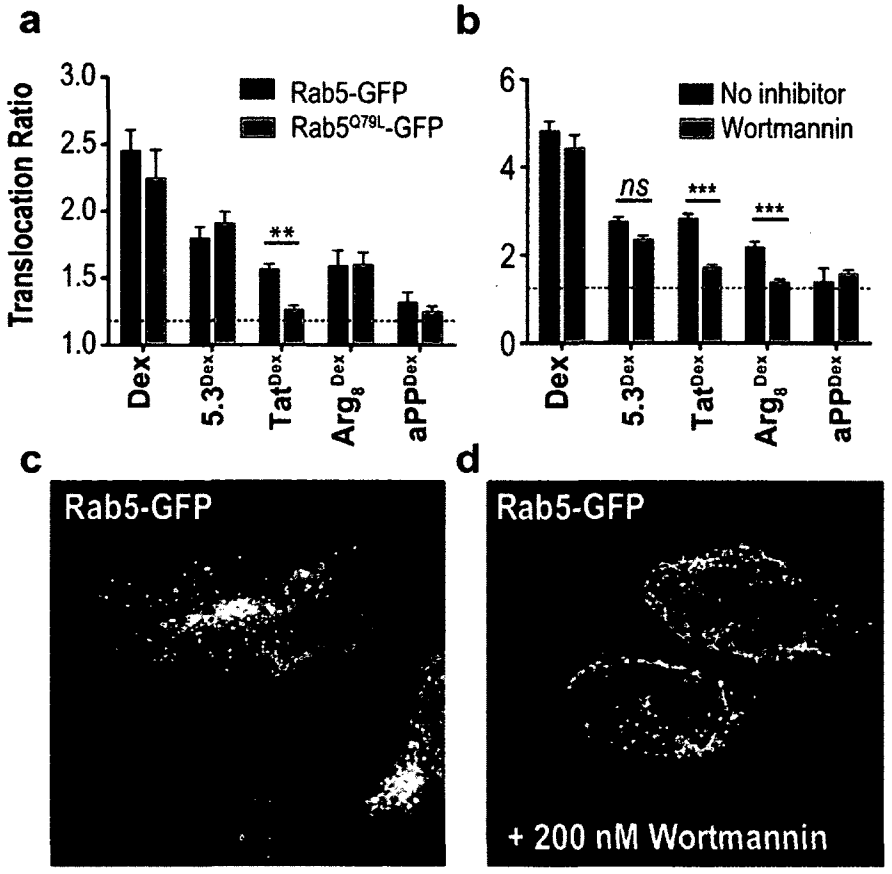


Figure 4.8: Cationic miniature protein **5.3** escapes from early endosomes. (**a**) Transfection with $Rab5^{Q79L}$-GFP does not block the increase of translocation ratio seen after treatment with dexamethasone, $\mathbf{5.3}^{Dex}$, or $\mathbf{Arg8}^{Dex}$, but decreases the translocation ratio measured after treatment with $\mathbf{Tat}^{Dex}$ ($p = 0.005$). (**b**) HeLa cells treated with 200 nM wortmannin for 30 min before treatment with 1 μM dexamethasone, $\mathbf{5.3}^{Dex}$, $\mathbf{Tat}^{Dex}$, $\mathbf{Arg8}^{Dex}$, or $\mathbf{aPP}^{Dex}$ for 30 additional min in the continued presence of the drug. Wortmannin decreased the translocation ratio measured after treatment with $\mathbf{Tat}^{Dex}$ ($p = 8.1 \times 10^{-15}$) and $\mathbf{Arg8}^{Dex}$ ($p = 2.3 \times 10^{-5}$). The effect of wortmannin on the distribution of Rab5 is assessed by confocal microscopy. (**c**) Rab5-GFP expressing HeLa cell. (**d**) Rab5-GFP expressing HeLa cell in the presence of 200 nM wortmannin for 30 min. *ns*, not significant; $*p \leq 0.05$; $**\ p \leq 0.01$; $***\ p \leq 0.001$, ANOVA with Bonferroni post test.

of Rab5+ vesicles[226, 224]. This concentration of wortmannin was sufficient to induce tubulation and enlargement of Rab5+ vesicles (Fig. 4.8d), observations consistent with the inhibition of their maturation. As can be seen in Fig. 4.8b, this treatment blocked the increase in translocation ratio seen after treatment with $\mathbf{Tat}^{Dex}$ and $\mathbf{Arg8}^{Dex}$ by 72% and 77%, respectively. In contrast, the translocation ratio seen after treatment with 5.3^{Dex} in the presence or absence of wortmannin was similar (2.36 ± 0.08 *vs* 2.76 ± 0.10), confirming the ability of 5.3^{Dex} to escape from Rab5+ vesicles. The difference between results obtained after wortmannin treatment and overexpression of $Rab5^{Q79L}$-GFP may be explained by recent results that wortmannin may arrest vesicles in an early Rab5+ state, prior to the recruitment of Rab5 effectors[226, 205]. Taken together, these data suggest that arresting vesicle maturation inhibits the cytosolic access of both $\mathbf{Tat}^{Dex}$ and $\mathbf{Arg8}^{Dex}$, but that 5.3^{Dex} is capable of leaving these earliest vesicles and reaching the cytoplasm.

4.2.4 Arginine topology affects protonation state transitions during endocytic acidification.

While the above data suggests that cationic miniature protein **5.3** can access the cytoplasm by escaping from Rab5+ vesicles, the chemical basis for why this sequence possesses this property is unknown. Comparisons between the sequence of **5.3**, a miniature protein that is cell penetrating, and the sequences of **aPP**, **4.2**, **4.3**, and **5.2**, miniature proteins which do not enter the cytoplasm, provide an opportunity to examine the effects of sequence changes on side-chain protonation states and structure. Data presented in figure 3.2 show that arginine containing variants of miniature proteins have similar circular dichroism spectra, suggesting a conserved fold. Large scale structural differences among these miniature proteins are thus unlikely to be the basis for their differing properties. Furthermore, the number of arginine residues alone, 5 versus 4, is not sufficient to explain the differences in cell permeability.

We wondered whether side-chain protonation states could be altered by changes in

neighboring amino acids, and whether these differences may explain differences in the ability to access to the cytoplasm. Importantly, for the guanidinium side chain group of arginine $pK_a \approx 12$, indicating that this side chain is nearly always protonated in the buffered (pH =7.4) aqueous environment surrounding the cell, and shifting to the more acidic environment of the endosome is unlikely to alter this state. Therefore we instead looked to see if there were acidic side chains whose pK_a was shifted such that in they would undergo a protonation state transition within the endosome.

To begin, we submitted the amino acid sequence of **aPP**, 4.2, 4.3, 5.2 and 5.3 to PEPSTR[232] an online structural prediction tool that returns predicted structures based on sequence using hidden Markov modeling. These structures were essentially identical to the X-ray structure of **aPP[129]** bolstering our expectation that these sequences fold in a similar manner. Repeating this structural prediction using another method based on neural networks (available online at Geno3D[233]) showed the ten lowest energy predicted conformations of 5.3 based on sequence had an overall RMSD of 0.49Å from **aPP**, providing further reassurange that the backbone fold is maintained between these two miniature proteins. We compared the sequences of **aPP** and 5.3 and identified residues that are altered in 5.3. These are Glu15, Asn22, Gln25, and Gln26, while Arg19 remains unchanged. Finally, we used the software tool PKATOOL[234], an algorithm that predicts amino acid pK_a based on three dimentional structure, to search for interactions that alter side-chain pK_a in **aPP** and 5.3 using the published X-ray structure of **aPP** and the predicted structure of 5.3 obtained from PEPSTR. First, we analyzed the side-chain interactions within **aPP** looking for interactions that would be destroyed by the arginine substitutions in 5.3. Second, we asked whether any new interactions would be created by the arginine subsitutions within 5.3. Third, drawing on our experimental evidence that blocking endosomal acidification blocks cytoplasmic access, we asked which, if any, residues within 5.3 would be predicted to undergo a protonation state transition from pH 7.4 (the pH of the extracellular medium) to pH 4 (the pH of the lysosome). This range is likely overly broad as our data supports a model in which 5.3 can escape from

the less acidic early endosome.

Our analysis (shown in Fig. 4.9) reveals that within **aPP**, Glu15 interacts strongly with both Asp16 ($kT/e = 1.003$) and Arg19 ($kT/e = -1.078$). In fact, these interaction energies are the second and third largest in magnitude within the molecule, and two of only three interactions with $|kT/e| > 1$. These residues are also positioned closely in space, with Glu15-Asp16 side chains within 2.8Å of each other, and Glu15-Arg19 distance to be ~3.2Å (Fig. 4.9a). Using pKaTool, the charge state of each of these amino acids is shown as a function of pH in figure 4.9b. Both Glu15 and Asp16 begin to undergo protonation as the pH drops below 7. One might suspect that the close apposition of the glutamic acid with the basic arginine stabilizes this amino acid in the deprotonated state, however when a third titratable species is added close by (Asp16) the overall effect is less obvious. Modeling this effect with the pKaTool shows a reciporical relationship in pKa purterbation. That is to say, that modeling an increased pKa of Asp16, results in a decreased pKa of Glu15, and vice versa (Fig. 4.9c). Because amino-acid side chains far from their equilibrium may not behave according to Henderson-Hasselbalch transitions, the concept of $pK_{1/2}$ is introduced as the pH at which the acid is half protonated. This value is similar to the value of the pK_a but does not imply knowledge of protonation state occupancy at other pH conditions, because, especially in protein, other acid-base transitions may be taking place. The $pK_{1/2}$ of Asp16 is predicted to be at pH = 4.2, changing Glu15 $\rightarrow$ Arg (i.e. modeling this residue as a base with pKa = 13, see figure 4.9d) changes this value to $pK_{1/2} = 3.4$, indicating stabilization of the deprotonated species. Further changing Asn22 $\rightarrow$ Arg (shown in gray) increases this difference, with the predicted $pK_{1/2} = 3.1$. Thus arginine substitutions at residues 15 and 22 is likely to stabilize Asp16 in the deprotonated state giving 5.3 a *decreased* positive charge at similar pH than would be present with an alternative spatial arrangement of charges.

Next we examined the modeled structure of 5.3 obtained from Pepstr. Interaction energies examined in pKaTool revealed no new interactions with energies $|kT/e| > 1$. The energies of interaction of Glu15-Asp16 and Glu15-Arg19 previously seen were now

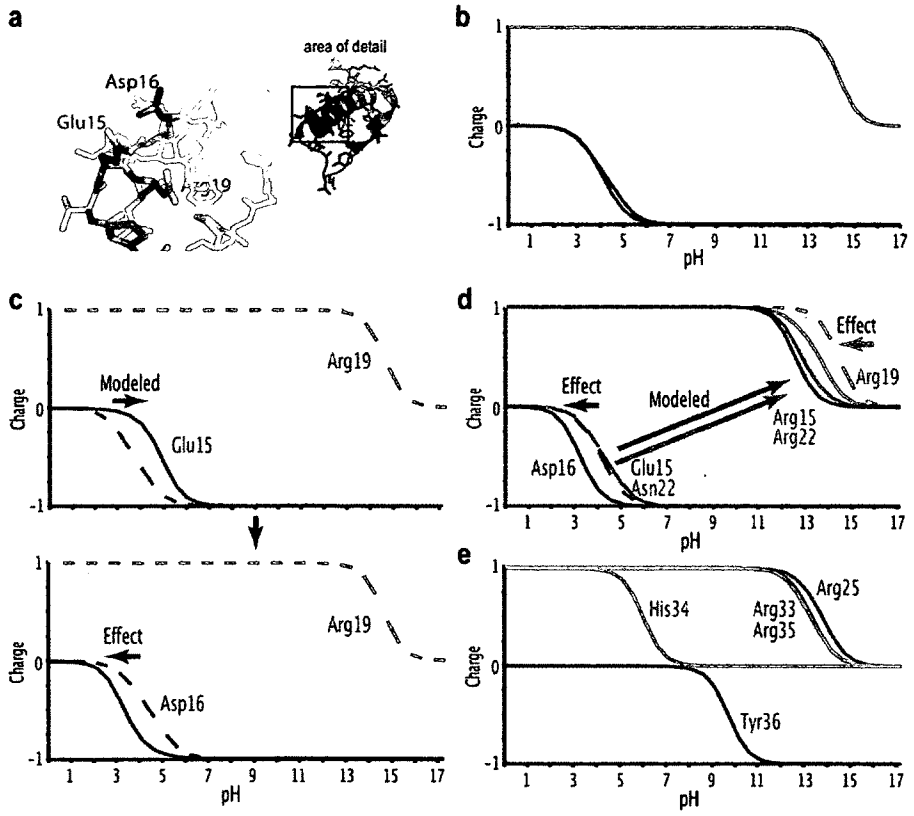


Figure 4.9: Modeling arginine interactions reveals subtly perturbed side-chain pK_a's, but these interactions stabilize deprotonated carboxylic acids, making protonation state transitions less likely to occur during endosomal acidification. (**a**) X-ray structure (PDB code 1PPT) showing Asp16, Glu15, and Arg19 are closely approximated in space. (**b**) Modeled *pH* titration of Asp16, Glu15, and Arg19 with color coding as shown in panel **a**, Asp16 and Glu15 become protonated (and neutrally charged) as the *pH* drops from 7 to 3. (**c**) The pK_a of Glu15 and Asp16 are linked. Glu15 as more basic (shift shown by green arrow) stabilizes the deprotonated form of Asp16 resulting in a lower modeled pK_a (blue arrow). The modeled pK_a of Arg19 remains unchanged. (**d**) Modeling Glu15 and Asn22 as arginine residues is predicted to increase the acidity of Asp16 and Arg19. (**e**) His34 is the only residue predicted to become protonated within the early endosome.

decreased bewteen both Arg15 and Asp16 ($kT/e = -0.252$) as well as Arg15 and Arg19 ($kT/e = 0.347$). Decreasing the threshold for significant interactions $|kT/e| > 0.5$ revealed a continued interaction between Asp23 and Asp16 whose strength is increased slightly from $kT/e = 0.402$ to $kT/e = 0.501$, as well as new interactions present between Arg25 and Tyr21 ($kT/e = -0.504$) as well as Arg26 and Asp23 ($kT/e = -0.578$). Searching for side-chains whose pK_a's were perturbed by molecular interactions revealed that residues Arg15, Asp16, Arg19, Asp23, and Arg25 showed the greatest deviation of $|\Delta pK_a| \geq 0.5$, while Tyr21 and Arg26 were perturbed less. Given that cationic miniature protein 5.3 can escape from the moderatly acidic early endosome, we searched for effects that would shift protonation state changes seen under acidic conditions to a more neutral pH range. Removing the influence of Arg19 lead to a decrease in the predicted acidity of Asp16 from $pK_{1/2} = 3.6$ to $pK_{1/2} = 4.2$ suggesting that Arg19 stabilizes the deprotonated form of Asp16. A similar effect is seen on Asp23, where removing Arg19 resulted in a predicted shift of $pK_{1/2} = 3.9$ to $pK_{1/2} = 4.2$. Removing either or both Arg25 and Arg26 also was predicted to destabilize Asp16 and Asp23, though the predicted effect is smaller. These results suggest that stabilization of the deprotonated forms of Asp16 or Asp23 is improved in the presence of the additional arginine residues, and that arginine leads to a more acidic transition rather than one that would take place at a more neutral pH.

Finally, we asked whether any side chains of 5.3 would be predicted to undergo signficant protonation state changes between $pH\,8$ and $pH\,5$, roughly the pH transition that takes place between the extracellular space and the late endosomal compartment. The only residue of 5.3 that is predicted to undergo a protonation state transition within this range is His34 with a predicted $pK_{1/2} = 5.9$, which transitions from a neutral to a positively charged species (Fig. 4.9e). The interaction energy prediction matrix shows that this residue possesses only minimal interactions with other residues. Removing either Arg15, 19, 22, 25, 26, or 33 resulted in increases in the calculated $pK_{1/2}$ from ~5.9 to 6.0-6.1. While this difference is small it is consistent with an intuitive model in which

the presence of each of additional arginines destabilizes the protonated form of the side chain. His34 in the context of **aPP** is predicted to have a $pK_{1/2} = 6.0$.

None of the above outlined predictions identify a new protonation state transition that is predicted to take place during acidification of the early endosome. This analysis views miniature proteins **aPP** and 5.3 in isolation, without taking into account known interactions with cellular proteins and proteoglycans. Based on this limited view, my analysis is unable to identify particular side chain protonation states that predicted to be altered between the extracellular $pH7$ environment and the and more acidic endosomal environment, any residue within 5.3 whose transition is predicted to take place in the early endosome where the analogus residue in **aPP** is not also predicted to do the same. Of course, this does not exclude the possibility that such a change does, in fact, take place – interactions between either **aPP** or 5.3 with cellular proteins could alter the individual pK_as of the peptide side chains, leading to the formation of a critical protonated species that is able to escape from the endosome. Alternatively, the activity of the endosomal V-type ATPase could be required for escape for a separate reason. The activity of V-type ATPase contributes to a trans-membrane electrochemical gradient, leads to the influx of chloride, across chloride channels, and participates in numerous other physiologic ion movements. One or several of these features could regulate the activity an endosomal protein required for the cytoplasmic escape of cationic miniature proteins.

4.3 Discussion

Coupling a new assay for cytosolic localization with live cell confocal microscopy allowed us to compare the structural determinants required for efficient endocytosis with those required for cytoplasmic escape. This comparison revealed that contrasting structural determinants control endocytic uptake and endosomal release. Uptake is favored by clustered α-helical arginine side chains, whereas release requires a more dispersed arginine array.

I found that inhibiting endocytosis blocked the ability of miniature protein 5.3, **Tat**, and **Arg8** to reach the cytoplasm. This finding strongly suggests against their ability to penetrate the plasma membrane directly, and argues for a model in which their pathway to the cytoplasm to begin with internalization into cellular vesicles followed by escape to the cytosol. I found that the internalization of 4.2, 4.3, 5.2, 5.3 as well as **Tat** and **Arg8** was dependent on the activity of dynamin, and was sensitive to cholesterol depletion. My finding that the small molecule EIPA blocked internalization suggests that reorganization of the actin cytoskeleton is also involved in the internalization of these miniature proteins and peptides. Internalization by pinocytic pathways such as the RhoA and Arf6 pathway and clathrin mediated endocytosis are consistent with these observations. Accordingly, after 30 minutes of internalization I observed the presence 5.3^R, $\mathbf{Tat}^R$ and $\mathbf{Arg8}^R$ within Rab5+ vesicles, followed by Rab7+ vesicles.

Previous studies on the trafficking of various viruses[31] and clostridial toxins, such as botulinum and anthrax toxin[32], have revealed multiple potential release points along the endocytic pathway. JC virus[229], SV40[196], and HIV-1[235], all require transport beyond early Rab5+ endosomes to gain cytosolic access. Requirements for cytosolic access have been extensively studied in the case of SV40: transport from Rab5+ to Rab7+ endosomes precedes transport to the golgi, and the endoplasmic reticulum where the virus crosses the endoplasmic reticulum membrane to access the cytosol. Botulinum and tetanus toxins, however, seem to require only low pH to insert into the membranes[236], and thus cytoplasmic access can occur in the absence of endosome maturation.

Inhibiting Rab5 vesicle maturation either through the addition of wortmannin or expression of a Rab5 dominant negative mutant had no effect on the ability of 5.3 to reach the cytoplasm, indicating that 5.3 enters into and efficiently escapes from early (Rab5+) endosomes. In contrast, **Tat** and **Arg8**, which are also present in Rab5+ vesicles, require Rab protein exchange or the activity of Vps34 (and possibly the subsequent recruitment of Rab5 effectors) in order to reach the cytoplasm. Inhibition of endosomal acidification through the addition of bafilomycin inhibited cytosolic access of 5.3, **Tat** and **Arg8**.

Thus, the pathway uncovered here for miniature protein 5.3 bears the greatest resemblance to that described for botulinum toxin and anthrax toxin, which also escape from early endosomes and require low pH to do so. The pathways followed by **Tat** and **Arg8** require downstream events or further endosomal acidification that is achieved after the early endosome. This distinction implies the existence of alternate signals favoring early versus late endosomal release. Identifying these signals and understanding their mechanistic basis may allow researchers to target molecules to a desired delivery pathway, increasing the rapidity and ease with which designed peptides may access the cytosol. This understanding may also shed light on the fundamental nature of the permeability barrier presented by lipid bilayers.

Miniature protein 5.3, with the lowest charge density, was released at the earliest common point along the endocytic pathway, whereas **Tat** and **Arg8**, which carry higher charge density, were released later. The differences in arginine/lysine number and orientation in 5.3, **Tat**, and **Arg8** can affect side chain pK_a, and thus overall charge at a given pH, as well as affinity for as-yet-unidentified cellular targets. It is possible that the distinct arginine array in 5.3 favors the formation of a critical protonated species in the moderately acidic environment of the early endosome (pH $\approx$ 6.5), whereas the lower pH present in lysosomes may be required to generate an equivalent state for **Tat** and **Arg8**. It is equally possible that the discrete arginine array in 5.3 represents an export signal for a cellular machinery that has yet been identified.

4.4 Experimental Methods

4.4.1 Cell Culture and Transfections.

HeLa cells (ATCC, Manassas, VA) were grown in T-75 culture flasks containing Dulbecco's Modified Essential Medium (DMEM, Gibco Cat. #11995-065) supplemented with 10% FBS, and 100 U/mL each Penicillin and Streptomycin. Transient transfections were performed using Fugene 6 or XtremeGene HP (Roche) using the protocols recommended

by the manufacturer. Plasmids encoding Rab5-GFP and Rab5Q79L-GFP were gifts from Pietro DeCamilli. A plasmid containing Rab7-GFP was a gift of Qing Zhong (Addgene plasmid # 28047). A plasmid containing GalT-EGFP was a gift of Jeniffer Lippincott-Schwartz (Addgene plasmid # 11937). A plasmid containing the glucocortcoid receptor fused to EGFP (pK7-GR-GFP) was a gift of Ian Macara (Addgene plasmid # 15534).

4.4.2 Colocalization of miniature proteins with Alexa-488-transferrin and Rab-GFP fusions.

To examine the colocalization of rhodamine labeled miniature proteins or peptides with Alexa-488-transferrin (Tf488), HeLa cells were plated (200 µL, 104 cells/well, 96 well glass bottom plates, Matrical) the day prior to experiments. The media was replaced with 150 µL Hepes-Krebs-Ringer's (HKR) buffer (140 mM NaCl, 2 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM Hepes pH 7.4) containing 1 µM miniature protein or 25 nM Alexa-488-transferrin (Molecular Probes) and the cells incubated at 37°C for 30 min. The cells were then rinsed twice with 200 µL HKR buffer and nuclei were labeled by overlaying 200 µL HKR containing 300 nM Hoechst 33342 (Molecular Probes Cat. # H3570) for 5 minutes. Images of cells were acquired using a PerkinElmer LiveView spinning disk confocal microscope fitted with a 60 x 1.2 NA objective. Colocalization with Rab-GFP fusions was examined in an analogous way using HeLa cells transfected with the appropriate expression plasmid.

Colocalization between rhodamine and fluorescein labeled miniature proteins and peptides, or between rhodamine labeled miniature proteins and markers of endocytosis was completed using confocal microscopy. Live cell confocal microscopy was conducted using a PerkinElmer LiveView spinning disk confocal microscope fitted with heated microscope enclosure (maintained at 37°C), a 60x 1.2 NA objective, laser illumination, a Nipkow CSU-1X spinning disk, and a Hammamatsu CS9100 EMCCD camera operated using Velocity software. Alexafluor-488-transferrin was detected using a 488 nm laser

and a 520 nm long pass filter; tetraethylrhodamine and Alexafluor-546-transferrin were detected using a 546 nm laser line and a 580 nm long pass filter, Hoescht 33342 was detected using a 405 nm laser line and a 450 nm long pass filter. Images were processed for presentation by using the 'subtract background' option and linear contrast adjustments in Image J. All adjustments were applied uniformly to the entire image. Colocalization analysis was performed using the JACOP plugin5, available on the ImageJ website. The fraction of rhodamine signal overlapping with EGFP signal ($R \in G$) was calculated using Mander's coefficient and automatic thresholding. Pearson's correlation coefficient (R) was also obtained using the JACOP plugin.

4.4.3 Effects of inhibitors on cell uptake.

HeLa cells grown for 24 h in glass bottom plates were incubated with HKR buffer, or HKR buffer containing 80 μM dynasore, 50 μM N-ethyl-isopropyl-amiloride, or 5 mM methyl-β-cyclodextrin for 30 min at 37°C prior to the addition of 1 μM rhodamine-labeled miniature protein or peptide (Tat and Arg8). The cells were washed twice with DMEM, the nuclei labeled with Hoescht, and images acquired as described above for colocalization experiments.

4.4.4 Effect of inhibitors on cytoplasmic access by dex-labeled miniature proteins or peptides.

HeLa cells transfected with pK7-GR-GFP for 24 - 36 hr were plated (104/well in 200 μL) the day prior to experiments in 96 well glass bottom plates (Matrical). The media was replaced with HKR buffer containing 300 nM Hoescht and the cells incubated for 30 min at 37°C. The media was then replaced with 150 μL HKR buffer or HKR buffer containing 1 μM dexamethasone labeled cationic miniature protein or peptides or 1 μM dexamethasone as a positive control. The cells were incubated for a further 30 min at 37°C before epifluorescence imaging using a Zeiss Axiovert 200M microscope fitted with

a mRM digital camera, a 63 x 1.3NA PlanApo objective and an EXFO-Excite illumination source. Dual color epifluorescence images were acquired using the following filter sets: Hoescht, Zeiss Filter Set #49 (ex G 365 nm, FT 395, em BP 445/50 nm); EGFP, Zeiss Filter Set #44 (ex BP 475/50 nm, FT 500, em BP 530/50 nm).

To examine the effects of various inhibitors on the ability of dexamethasone labeled cationic miniature proteins to activate the glucocorticoid receptor, HeLa cells were transfected and plated as above. Thirty minutes prior to the addition of dexamethasone labeled cationic miniature proteins and peptides the media was replaced with HKR buffer containing 300 nM Hoescht and 80 µM dynasore, 50 µM N-ethyl-isopropyl-amiloride or 5 mM methyl-β-cyclodextrin and the cells incubated for 30 min at 37°C after which 1 µM dexamethasone labeled miniature protein or peptide was added. Uptake was allowed to proceed for an additional 30 min at 37°C before analysis by epifluorescence images as described above. Images were analyzed to determine the ratio of nuclear to cytoplasmic GR-GFP using CellProfiler (see § 3.4.5 on page 84).

4.4.5 pGR-mCherry vector construction.

The full length glucocorticoid receptor (NR3C1) was subcloned into pmCherry-N1 (Invitrogen), generating a C-terminal fluorescent protein fusion, by amplifying the gene from pK7-GR-GFP using PCR primers containing Sac1 and BamH1 restriction sites. The amplified PCR product was digested and ligated into the multiple cloning site of pmCherry-N1 digested with the same restriction enzymes.

4.4.6 Requirement of Rab5 activity for the ability of peptides to reach the cytoplasm.

HeLa cells transfected for 24 - 36 hr with pGR-mCherry and either Rab5-GFP or Rab5^{Q79L}-GFP were plated (10⁴/well in 200 µL) the day prior to experiments in 96 well glass bottom plates (Matrical). The cells were washed with HKR buffer and the nuclei labeled

using Hoescht as above. Then 150 µL HKR buffer or HKR buffer containing 1 µM dexamethasone labeled cationic miniature protein or peptides or 1 µM dexamethasone as a positive control was overlaid into the well. The cells were incubated for a further 30 min at 37°C before imaging. Cells that had been cotransfected with EGFP fusions were identified via their characteristic pattern of green and red fluorescence. To evaluate the extent of GR activation, epifluorescence images were obtained as described above, except the fluorescence of pGR-mCherry was also visualized using Zeiss Filter Set #43 (ex BP 545/25 nm, FT 570, em BP 605/70 nm). Images were analyzed to determine the ratio of nuclear to cytoplasmic GR-mCherry using CellProfiler.

4.4.7 Statistical Analysis.

Comparisons within groups were made using ANOVA. Pairwise comparisons within groups were made using Bonferroni's post-test after finding a significant difference using ANOVA. P-values are adjusted for the number of comparisons (using Bonferroni's method[237, 238, 239]) so that the family-wise error rate, FWER = 0.05 as follows. Following the determination of a statistically significant difference using ANOVA, the significance of a given pairwise comparison from among set containing J groups is calculated using the t-stastistic. Let $p(t, \mathrm{df})$ be the significance of a two-tailed t-test between two groups with degrees of freedom df. Then the signiciance of the difference using Bonferroni's method p' between two groups from a family containing J groups is distributed as

$$p' = p(t, N - J) \cdot \frac{J(J-1)}{2}$$

where t is the test statistic from the between group comparison, N is the number of degrees of freedom from the ANOVA analysis and J is the number of groups in the family. A critical feature of this method of correction is that the number of groups in the family from which the pairwise comparison is drawn has a significant effect on the

the significance of the result, p'. An alternative way of reporting Bonferroni corrected P-values is to alter the significance level α at which a test is considered signficant. In this case, α is usually set so that $\alpha' = \alpha \cdot \frac{2}{J \cdot (J-1)}$, and the null hypothesis H^0 is rejected if $p(t, N-J) > \alpha'$. As an example using this second method, a single pairwise comparison from a family containing 6 members would need $p < 0.003$ to be considered significant. This alternative method is functionally equivalent to the one used here. The advantage of reporting p' is that these values are intuitively understandable and the false-discovery rate (in the alternative method, the example comparison with $p < 0.003$ has a false discovery rate of 0.05). One caveat to our approach is that in the case of negative results (i.e. when H^0 is not rejected), it is possible to find $p' > 1$, which is unphysical. This oddity is generally acceptable, because in these cases, the exact value of the significance of the test in unimportant.

Bibliography

[1] O P Hamill, A Marty, E Neher, B Sakmann, and F J Sigworth. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflugers Arch.*, 391(2):85–100, August 1981.

[2] C A Lipinski, F Lombardo, B W Dominy, and P J Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*, 46(1-3):3–26, March 2001.

[3] Yanay Ofran and Burkhard Rost. Protein-protein interaction hotspots carved into sequences. *PLoS Comput. Biol.*, 3(7):e119, July 2007.

[4] Warren L DeLano. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.*, 12(1):14–20, February 2002.

[5] T Clackson and J A Wells. A hot spot of binding energy in a hormone-receptor interface. *Science*, 267(5196):383–386, January 1995.

[6] Ozlem Keskin, Buyong Ma, and Ruth Nussinov. Hot regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. *J Mol Biol*, 345(5):1281–1294, February 2005.

[7] Thomas A Edwards and Andrew J Wilson. Helix-mediated protein–protein interactions as targets for intervention using foldamers. *Amino Acids*, 41(3):743–754, August 2011.

[8] Benzhuo Lu and Chung F Wong. Direct estimation of entropy loss due to reduced translational and rotational motions upon molecular binding. *Biopolymers*, 79(5): 277–285, December 2005.

[9] Amit R Reddi, Tabitha R Guzman, Robert M Breece, David L Tierney, and Brian R Gibney. Deducing the energetic cost of protein folding in zinc finger proteins using designed metallopeptides. *J Am Chem Soc*, 129(42):12815–12827, October 2007.

[10] Thomas A Knotts, Nitin Rathore, and Juan J de Pablo. An entropic perspective of protein stability on surfaces. *Biophys J*, 94(11):4473–4483, June 2008.

[11] Francesco Colizzi, Remo Perozzo, Leonardo Scapozza, Maurizio Recanatini, and Andrea Cavalli. Single-Molecule Pulling Simulations Can Discern Active from Inactive Enzyme Inhibitors. *J Am Chem Soc*, 132(21):7361–7371, June 2010.

[12] William L Jorgensen. Efficient drug lead discovery and optimization. *Acc. Chem. Res.*, 42(6):724–733, June 2009.

[13] William L Jorgensen. Drug discovery: Pulled from a protein's embrace. *Nature*, 466(7302):42–43, July 2010.

[14] Jason W Chin and Alanna Schepartz. Design and Evolution of a Miniature Bcl-2 Binding Protein . *Angew. Chem. Int. Ed. Engl.*, 40(20):3806–3809, October 2001.

[15] Neal J Zondlo and Alanna Schepartz. Highly Specific DNA Recognition by a Designed Miniature Protein. *J Am Chem Soc*, 121(29):6938–6939, July 1999.

[16] Mary Kay H Pflum. Grafting Miniature DNA Binding Proteins. *Chem Biol*, 11(1): 3–4, January 2004.

[17] Crystal D Zellefrow, Jennifer S Griffiths, Sarmistha Saha, Abby M Hodges, Jessica L Goodman, Joshiawa Paulk, Joshua A Kritzer, and Alanna Schepartz. Encodable activators of SRC family kinases. *J Am Chem Soc*, 128(51):16506–16507, December 2006.

[18] Anja C Gemperli, Stacey E Rutledge, Abby Maranda, and Alanna Schepartz. Paralog-selective ligands for bcl-2 proteins. *J Am Chem Soc*, 127(6):1596–1597, February 2005.

[19] Dasantila Golemi-Kotra, Rachel Mahaffy, Matthew J Footer, Jennifer H Holtzman, Thomas D Pollard, Julie A Theriot, and Alanna Schepartz. High affinity, paralog-specific recognition of the Mena EVH1 domain by a miniature protein. *J Am Chem Soc*, 126(1):4–5, January 2004.

[20] Jennifer H Holtzman, Kamil Woronowicz, Dasantila Golemi-Kotra, and Alanna Schepartz. Miniature protein ligands for EVH1 domains: interplay between affinity, specificity, and cell motility. *Biochemistry*, 46(47):13541–13553, November 2007.

[21] Joshua A Kritzer, Reena Zutshi, Mingtatt Cheah, F Ann Ran, Rachel Webman, Taritree M Wongjirad, and Alanna Schepartz. Miniature protein inhibitors of the p53-hDM2 interaction. *Chembiochem*, 7(1):29–31, 2006.

[22] Jin Kim Montclare and Alanna Schepartz. Miniature homeodomains: high specificity without an N-terminal arm. *J Am Chem Soc*, 125(12):3416–3417, March 2003.

[23] Arjel D Bautista, Cody J Craig, Elizabeth A Harker, and Alanna Schepartz. Sophistication of foldamer form and function in vitro and in vivo. *Current Opinion in Chemical Biology*, 11(6):685–692, December 2007.

[24] E Harker and A Schepartz. Cell-Permeable b-Peptide Inhibitors of p53/hDM2 Complexation. *Chembiochem*, 2009.

[25] Joshua A Kritzer, James D Lear, Michael E Hodsdon, and Alanna Schepartz. Helical beta-peptide inhibitors of the p53-hDM2 interaction. *J Am Chem Soc*, 126(31):9468–9469, August 2004.

[26] Joshua A Kritzer, Michael E Hodsdon, and Alanna Schepartz. Solution structure of a beta-peptide ligand for hDM2. *J Am Chem Soc*, 127(12):4118–4119, March 2005.

[27] Joshua A Kritzer, Olen M Stephens, Danielle A Guarracino, Samuel K Reznik, and Alanna Schepartz. beta-Peptides as inhibitors of protein-protein interactions. *Bioorg Med Chem*, 13(1):11–16, January 2005.

[28] Olen M Stephens, Sunghwan Kim, Brett D Welch, Michael E Hodsdon, Michael S Kay, and Alanna Schepartz. Inhibiting HIV fusion with a beta-peptide foldamer. *J Am Chem Soc*, 127(38):13126–13127, September 2005.

[29] J W Chin, R M Grotzfeld, M A Fabian, and A Schepartz. Methodology for optimizing functional miniature proteins based on avian pancreatic polypeptide using phage display. *Bioorg Med Chem Lett*, 11(12):1501–1505, June 2001.

[30] Joshua A Kritzer, Nathan W Luedtke, Elizabeth A Harker, and Alanna Schepartz. A rapid library screen for tailoring beta-peptide structure and function. *J Am Chem Soc*, 127(42):14584–14585, October 2005.

[31] Billy Tsai. Penetration of nonenveloped viruses into the cytoplasm. *Annu Rev Cell Dev Biol*, 23:23–43, 2007.

[32] Lance L Simpson. Identification of the major steps in botulinum toxin action. *Annu. Rev. Pharmacol. Toxicol.*, 44:167–193, 2004.

[33] K Sandvig and B van Deurs. Entry of ricin and Shiga toxin into cells: molecular mechanisms and medical perspectives. *EMBO JOURNAL*, 19(22):5943–5950, November 2000.

[34] A Fontaine, J Arondel, and P J Sansonetti. Role of Shiga toxin in the pathogenesis of bacillary dysentery, studied by using a Tox- mutant of Shigella dysenteriae 1. *Infect Immun*, 56(12):3099–3109, December 1988.

[35] Anne L Ackerman, Christoph Kyritsis, Robert Tampé, and Peter Cresswell. Access of soluble antigens to the endoplasmic reticulum can explain cross-presentation by dendritic cells. *Nat Immunol*, 6(1):107–113, January 2005.

[36] Lianjun Shen and Kenneth L Rock. Priming of T cells by exogenous antigen cross-presented on MHC class I molecules. *Curr Opin Immunol*, 18(1):85–91, February 2006.

[37] P Cresswell. Intracellular surveillance: controlling the assembly of MHC class I-peptide complexes. *Traffic*, 1(4):301–305, April 2000.

[38] Peter Cresswell, Anne L Ackerman, Alessandra Giodini, David R Peaper, and Pamela A Wearsch. Mechanisms of MHC class I-restricted antigen processing and cross-presentation. *Immunol. Rev.*, 207:145–157, October 2005.

[39] R Steinman. Dendritic Cells In Vivo: A Key Target for a New Vaccine Science. *Immunity*, 29(3):319–324, September 2008.

[40] J W Yewdell, C C Norbury, and J R Bennink. Mechanisms of exogenous antigen presentation by MHC class I molecules in vitro and in vivo: implications for generating CD8+ T cell responses to infectious agents, tumors, transplants, and vaccines. *Adv. Immunol.*, 73:1–77, 1999.

[41] Vasso Apostolopoulos, Dodie S Pouniotis, Peter J van Maanen, Rene W Andriessen, Jodie Lodding, Pei-Xiang Xing, Ian F C McKenzie, Bruce E Loveland, and Geoffrey A Pietersz. Delivery of tumor associated antigens to antigen presenting cells using penetratin induces potent immune responses. *Vaccine*, 24(16):3191–3202, April 2006.

[42] K Simons and D Toomre. Lipid rafts and signal transduction. *Nat Rev Mol Cell Biol*, 1(1):31–39, October 2000.

[43] P P Di Fiore and P De Camilli. Endocytosis and signaling. an inseparable partnership. *Cell*, 106(1):1–4, July 2001.

[44] M Kovacsovics-Bankowski and K Clark. Efficient major histocompatibility complex class I presentation of exogenous antigen upon phagocytosis by macrophages. In *Proceedings of the ...*, 1993.

[45] F Carbone. Class I-restricted processing and presentation of exogenous cell-associated antigen in vivo. *Journal of Experimental Medicine*, 1990.

[46] M J Bevan. Cross-priming for a secondary cytotoxic response to minor H antigens with H-2 congenic cells which do not cross-react in the cytotoxic assay. *J Exp Med*, 143(5):1283, 1976.

[47] Tom A M Groothuis and Jacques Neefjes. The many roads to cross-presentation. *J Exp Med*, 202(10):1313–1318, November 2005.

[48] Lianjun Shen, Luis J Sigal, Marianne Boes, and Kenneth L Rock. Important role of cathepsin S in generating peptides for TAP-independent MHC class I crosspresentation in vivo. *Immunity*, 21(2):155–165, August 2004.

[49] M Grommé, F G Uytdehaag, H Janssen, J Calafat, R S van Binnendijk, M J Kenter, A Tulp, D Verwoerd, and J Neefjes. Recycling MHC class I molecules and endosomal peptide loading. *Proc Natl Acad Sci USA*, 96(18):10326–10331, August 1999.

[50] Robert J Binder and Pramod K Srivastava. Peptides chaperoned by heat-shock proteins are a necessary and sufficient source of antigen in the cross-priming of CD8+ T cells. *Nat Immunol*, 6(6):593–599, June 2005.

[51] Joost Neijssen, Carla Herberts, Jan Wouter Drijfhout, Eric Reits, Lennert Janssen, and Jacques Neefjes. Cross-presentation by intercellular peptide transfer through gap junctions. *Nature*, 434(7029):83–88, March 2005.

[52] B Koppelman, J Blum, M Marks, and P Cresswell. Co-localization of molecules involved in antigen processing and presentation in an early endocytic .... *Nature*, 1990.

[53] Galit Denkberg and Yoram Reiter. Recombinant antibodies with T-cell receptor-like specificity: novel tools to study MHC class I presentation. *Autoimmun Rev*, 5 (4):252–257, April 2006.

[54] Galit Denkberg, Cyril J Cohen, Avital Lev, Patrick Chames, Hennie R Hoogenboom, and Yoram Reiter. Direct visualization of distinct T cell epitopes derived from a melanoma tumor-associated antigen by using human recombinant antibodies with MHC- restricted T cell receptor-like specificity. *Proc Natl Acad Sci USA*, 99 (14):9421–9426, July 2002.

[55] A Porgador, J W Yewdell, Y Deng, J R Bennink, and R N Germain. Localization, quantitation, and in situ detection of specific peptide-MHC class I complexes using a monoclonal antibody. *Immunity*, 6(6):715–726, June 1997.

[56] B Berwin and C V Nicchitta. To find the road traveled to tumor immunity: the trafficking itineraries of molecular chaperones in antigen-presenting cells. *Traffic*, 2 (10):690–697, October 2001.

[57] Marta Fernández-Suárez and Alice Y Ting. Fluorescent probes for super-resolution imaging in living cells. *Nat Rev Mol Cell Biol*, 9(12):929–943, December 2008.

[58] J Fölling, V Belov, R Kunetsky, R Medda, A Schönle, A Egner, C Eggeling, M Bossi, and S W Hell. Photochromic Rhodamines Provide Nanoscopy with Optical Sectioning. *Angew. Chem. Int. Ed.*, 46(33):6266–6270, August 2007.

[59] Rajesh Babu Sekar and Ammasi Periasamy. Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *Journal of Cell Biology*, 160(5):629–633, March 2003.

[60] Melissa A Lowder, Jacob S Appelbaum, Elissa M Hobert, and Alanna Schepartz. Visualizing protein partnerships in living cells and organisms. *Current Opinion in Chemical Biology*, 15(6):781–788, December 2011.

[61] Prasanna Venkatraman, Tina T Nguyen, Matthieu Sainlos, Osman Bilsel, Sriram Chitta, Barbara Imperiali, and Lawrence J Stern. Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells. *Nat Chem Biol*, 3(4):222–228, April 2007.

[62] Nathan W Luedtke, Rachel J Dexter, Daniel B Fried, and Alanna Schepartz. Surveying polypeptide and protein domain conformation and association with FlAsH and ReAsH. *Nat Chem Biol*, 3(12):779–784, December 2007.

[63] Guido Gaietta, Thomas J Deerinck, Stephen R Adams, James Bouwer, Oded Tour, Dale W Laird, Gina E Sosinsky, Roger Y Tsien, and Mark H Ellisman. Multicolor and electron microscopic imaging of connexin trafficking. *Science*, 296(5567):503–507, April 2002.

[64] Guido M Gaietta, Ben N G Giepmans, Thomas J Deerinck, W Bryan Smith, Lucy Ngan, Juan Llopis, Stephen R Adams, Roger Y Tsien, and Mark H Ellisman. Golgi twins in late mitosis revealed by genetically encoded tags for live cell imaging and correlated electron microscopy. *Proc Natl Acad Sci USA*, 103(47):17777–17782, November 2006.

[65] Rachel J Dexter and Alanna Schepartz. Direct visualization of protein association in living cells with complex-edited electron microscopy. *Angew. Chem. Int. Ed. Engl.*, 49(43):7952–7954, October 2010.

[66] Stephen R Adams, Robert E Campbell, Larry A Gross, Brent R Martin, Grant K Walkup, Yong Yao, Juan Llopis, and Roger Y Tsien. New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J Am Chem Soc*, 124(21):6063–6076, May 2002.

[67] B A Griffin, S R Adams, and R Y Tsien. Specific covalent labeling of recombinant protein molecules inside live cells. *Science*, 281(5374):269–272, July 1998.

[68] Brent R Martin, Ben N G Giepmans, Stephen R Adams, and Roger Y Tsien. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat Biotechnol*, 23(10):1308–1314, October 2005.

[69] Fatemeh Madani, Jesper Lind, Peter Damberg, Stephen R Adams, Roger Y Tsien, and Astrid O Gräslund. Hairpin structure of a biarsenical-tetracysteine motif determined by NMR spectroscopy. *J Am Chem Soc*, 131(13):4613–4615, April 2009.

[70] Jessica L Goodman, E James Petersson, Douglas S Daniels, Jade X Qiu, and Alanna Schepartz. Biophysical and structural characterization of a robust octameric beta-peptide bundle. *J Am Chem Soc*, 129(47):14746–14751, November 2007.

[71] P J Bjorkman, M A Saper, B Samraoui, W S Bennett, J L Strominger, and D C Wiley. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature*, 329(6139):512–518, September 1987.

[72] P J Bjorkman, M A Saper, B Samraoui, W S Bennett, J L Strominger, and D C Wiley. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*, 329 (6139):506–512, September 1987.

[73] Rico Buchli, Rodney S Vangundy, Heather D Hickman-Miller, Christopher F Giberson, Wilfried Bardet, and William H Hildebrand. Real-time measurement of in vitro peptide binding to soluble HLA-A*0201 by fluorescence polarization. *Biochemistry*, 43(46):14852–14863, November 2004.

[74] H W Nijman, J G Houbiers, M P Vierboom, S H van der Burg, J W Drijfhout, J D'Amaro, P Kenemans, C J Melief, and W M Kast. Identification of peptide sequences that potentially trigger HLA-A2.1-restricted cytotoxic T lymphocytes. *Eur J Immunol*, 23(6):1215–1219, June 1993.

[75] Randi Vita, Laura Zarebski, Jason A Greenbaum, Hussein Emami, Ilka Hoof, Nima Salimi, Rohini Damle, Alessandro Sette, and Bjoern Peters. The immune epitope database 2.0. *Nucleic Acids Research*, 38(Database issue):D854–62, January 2010.

[76] P Sliz, O Michielin, J C Cerottini, I Luescher, P Romero, M Karplus, and D C Wiley. Crystal structures of two closely related but antigenically distinct HLA-A2/melanocyte-melanoma tumor-antigen peptide complexes. *J Immunol*, 167(6): 3276–3284, September 2001.

[77] Pamela A Wearsch and Peter Cresswell. Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer. *Nat Immunol*, 8(8):873–881, August 2007.

[78] David R Peaper and Peter Cresswell. The redox activity of ERp57 is not essential for its functions in MHC class I peptide loading. *Proc Natl Acad Sci USA*, 105(30): 10477–10482, July 2008.

[79] A Townsend, T Elliott, V Cerundolo, L Foster, B Barber, and A Tse. Assembly of MHC class I molecules analyzed in vitro. *Cell*, 62(2):285–295, July 1990.

[80] E J Baas, H M van Santen, M J Kleijmeer, H J Geuze, P J Peters, and H L Ploegh. Peptide-induced stabilization and intracellular localization of empty HLA class I complexes. *J Exp Med*, 176(1):147–156, July 1992.

[81] N A Hosken and M J Bevan. An endogenous antigenic peptide bypasses the class I antigen presentation defect in RMA-S. *J Exp Med*, 175(3):719–729, March 1992.

[82] K C Parker, B M Carreno, L Sestak, U Utz, W E Biddison, and J E Coligan. Peptide binding to HLA-A2 and HLA-B27 isolated from Escherichia coli. Reconstitution of HLA-A2 and HLA-B27 heavy chain/beta 2-microglobulin complexes requires specific peptides. *J Biol Chem*, 267(8):5451–5459, March 1992.

[83] S H Hsu, B Z Schacter, N L Delaney, T B Miller, V A McKusick, R H Kennett, J G Bodmer, D Young, and W F Bodmer. Genetic characteristics of the HeLa cell. *Science*, 191(4225):392–394, January 1976.

[84] K T Hogan and S L Brown. Localization and characterization of serologic epitopes on HLA-A2. *Hum Immunol*, 33(3):185–192, March 1992.

[85] B A Griffin, S R Adams, J Jones, and R Y Tsien. Fluorescent labeling of recombinant proteins in living cells with FlAsH. *Methods in Enzymology*, 327:565–578, 2000.

[86] R B Freedman. Protein disulfide isomerase: multiple roles in the modification of nascent secretory proteins. *Cell*, 57(7):1069–1072, June 1989.

[87] I Braakman, J Helenius, and A Helenius. Manipulating disulfide bond formation and protein folding in the endoplasmic reticulum. *EMBO JOURNAL*, 11(5):1717–1722, May 1992.

[88] T N Schumacher, M T Heemels, J J Neefjes, W M Kast, C J Melief, and H L Ploegh. Direct binding of peptide to empty MHC class I molecules on intact cells and in vitro. *Cell*, 62(3):563–567, August 1990.

[89] B Seliger, U Ritz, R Abele, M Bock, R Tampe, G Sutter, I Drexler, C Huber, and S Ferrone. Immune escape of melanoma: first evidence of structural alterations in two distinct components of the MHC class I antigen processing pathway. *Cancer Res*, 61(24):8647–8650, December 2001.

[90] David Stepensky, Naveen Bangia, and Peter Cresswell. Aggregate formation by ERp57-deficient MHC class I peptide-loading complexes. *Traffic*, 8(11):1530–1542, November 2007.

[91] B. Sadasivan, P.J. Lehner, B Ortmann, T. Spies, and P Cresswell. Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity*, 5(2):103–114, August 1996.

[92] Marco A Purbhoo, Darrell J Irvine, Johannes B Huppa, and Mark M Davis. T cell killing does not require the formation of a stable mature immunological synapse. *Nat Immunol*, 5(5):524–530, May 2004.

[93] Darrell J Irvine, Marco A Purbhoo, Michelle Krogsgaard, and Mark M Davis. Direct observation of ligand recognition by T cells. *Nature*, 419(6909):845–849, October 2002.

[94] S.A. Soper, E.B. Shera, J.C. Martin, J.H. Jett, J.H. Hahn, H.L. Nutter, and R.A. Keller. Single-molecule detection of rhodamine 6G in ethanolic solutions using continuous wave laser excitation. *Analytical chemistry*, 63(5):432–437, 1991.

[95] Pierre Guermonprez, Loredana Saveanu, Monique Kleijmeer, Jean Davoust, Peter Van Endert, and Sebastian Amigorena. ER-phagosome fusion defines an MHC class I cross-presentation compartment in dendritic cells. *Nature*, 425(6956):397–402, September 2003.

[96] Mathieu Houde, Sylvie Bertholet, Etienne Gagnon, Sylvain Brunet, Guillaume Goyette, Annie Laplante, Michael F Princiotta, Pierre Thibault, David Sacks, and Michel Desjardins. Phagosomes are competent organelles for antigen cross-presentation. *Nature*, 425(6956):402–406, September 2003.

[97] Etienne Gagnon, Sophie Duclos, Christiane Rondeau, Eric Chevet, Pamela H Cameron, Olivia Steele-Mortimer, Jacques Paiement, John J M Bergeron, and Michel Desjardins. Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. *Cell*, 110(1):119–131, July 2002.

[98] Anne L Ackerman, Alessandra Giodini, and Peter Cresswell. A role for the endoplasmic reticulum protein retrotranslocation machinery during crosspresentation by dendritic cells. *Immunity*, 25(4):607–617, October 2006.

[99] Anne L Ackerman, Christoph Kyritsis, Robert Tampé, and Peter Cresswell. Early phagosomes in dendritic cells form a cellular compartment sufficient for cross presentation of exogenous antigens. *Proc Natl Acad Sci USA*, 100(22):12889–12894, October 2003.

[100] N Touret, P Paroutis, M Terebiznik, R Harrison, S Trombetta, M Pypaert, A Chow, A Jiang, J Shaw, C Yip, I Mellman, and S Grinstein. Quantitative and Dynamic Assessment of the Contribution of the ER to Phagosome Formation. *Cell*, 123(1): 157–170, October 2005.

[101] Matthias Zehner, Achmet Imam Chasan, Verena Schuette, Maria Embgenbroich, Thomas Quast, Waldemar Kolanus, and Sven Burgdorf. Mannose receptor polyubiquitination regulates endosomal recruitment of p97 and cytosolic antigen translocation for cross-presentation. *Proceedings of the National Academy of Sciences*, 108(24): 9933–9938, June 2011.

[102] Jessica L Goodman, Daniel B Fried, and Alanna Schepartz. Bipartite tetracysteine display requires site flexibility for ReAsH coordination. *Chembiochem*, 10(10):1644–1647, July 2009.

[103] David A Armbruster and Terry Pry. Limit of blank, limit of detection and limit of quantitation. *Clin Biochem Rev*, 29 Suppl 1:S49–52, August 2008.

[104] A L Stout and D Axelrod. Evanescent field excitation of fluorescence by epi-illumination microscopy. *Appl Opt*, 28(24):5237–5242, December 1989.

[105] William R Strohl and David M Knight. Discovery and development of biopharmaceuticals: current issues. *Curr Opin Biotechnol*, 20(6):668–672, December 2009.

[106] Peter M Fischer. Cellular uptake mechanisms and potential therapeutic utility of peptidic cell delivery vectors: progress 2001-2006. *Med Res Rev*, 27(6):755–795, November 2007.

[107] Randolph M Johnson, Stephen D Harrison, and Derek Maclean. Therapeutic applications of cell-penetrating peptides. *Methods Mol Biol*, 683:535–551, 2011.

[108] Nathan W Luedtke, Peter Carmichael, and Yitzhak Tor. Cellular uptake of aminoglycosides, guanidinoglycosides, and poly-arginine. *J Am Chem Soc*, 125(41):12374–12375, October 2003.

[109] H RYSER, J B CAULFIELD, and J C AUB. Studies on protein uptake by isolated tumor cells. I. Electron microscopic evidence of ferritin uptake by Ehrlich ascites tumor cells. *J Cell Biol*, 14:255–268, August 1962.

[110] H RYSER, J C AUB, and J B CAULFIELD. Studies on protein uptake by isolated tumor cells. II. Quantitative data on the adsorption and uptake of I-131-serum albumin by Ehrlich ascites tumor cells. *J Cell Biol*, 15:437–449, December 1962.

[111] H J Ryser and R Hancock. Histones and basic polyamino acids stimulate the uptake of albumin by tumor cells in culture. *Science*, 150(695):501–503, October 1965.

[112] Hongyan Zhou, Shili Wu, Jin Young Joo, Saiyong Zhu, Dong Wook Han, Tongxiang Lin, Sunia Trauger, Geoffery Bien, Susan Yao, Yong Zhu, Gary Siuzdak, Hans R Schöler, Lingxun Duan, and Sheng Ding. Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell*, 4(5):381–384, May 2009.

[113] A D Frankel and C O Pabo. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, 55(6):1189–1193, December 1988.

[114] D Derossi, A H Joliot, G Chassaing, and A Prochiantz. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem*, 269(14):10444–10450, April 1994.

[115] D Derossi, S Calvet, A Trembleau, A Brunissen, G Chassaing, and A Prochiantz. Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. *J Biochem*, 271(30):18188–18193, July 1996.

[116] Jehangir S Wadia, Radu V Stan, and Steven F Dowdy. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat Med*, 10(3):310–315, March 2004.

[117] S R Schwarze, A Ho, A Vocero-Akbani, and S F Dowdy. In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science*, 285(5433): 1569–1572, September 1999.

[118] James J Cronican, David B Thompson, Kevin T Beier, Brian R Mcnaughton, Constance L Cepko, and David R Liu. Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *Acs Chem Biol*, 5(8):747–752, August 2010.

[119] James J Cronican, Kevin T Beier, Tina N Davis, Jen-Chieh Tseng, Weida Li, David B Thompson, Allen F Shih, Erin M May, Constance L Cepko, Andrew L Kung, Qiao Zhou, and David R Liu. A Class of Human Proteins that Deliver Functional Proteins into Mammalian Cells In Vitro and In Vivo. *Chem Biol*, 18(7): 833–838, July 2011.

[120] Jonathan B Rothbard, Theodore C Jessop, Richard S Lewis, Bryce A Murray, and Paul A Wender. Role of membrane potential and hydrogen bonding in the mechanism of translocation of guanidinium-rich peptides into cells. *J Am Chem Soc*, 126 (31):9506–9507, August 2004.

[121] Jonathan B Rothbard, Theodore C Jessop, and Paul A Wender. Adaptive translocation: the role of hydrogen bonding and membrane potential in the uptake of guanidinium-rich transporters into cells. *Adv Drug Deliv Rev*, 57(4):495–504, February 2005.

[122] Rainer Fischer, Mariola Fotin-Mleczek, Hansjörg Hufnagel, and Roland Brock. Break on through to the other side-biophysics and cell biology shed light on cell-penetrating peptides. *Chembiochem*, 6(12):2126–2142, December 2005.

[123] Rainer Fischer, Karsten Köhler, Mariola Fotin-Mleczek, and Roland Brock. A stepwise dissection of the intracellular fate of cationic cell-penetrating peptides. *J Biol Chem*, 279(13):12625–12635, March 2004.

[124] G Ter-Avetisyan, G Tunnemann, D Nowak, M Nitschke, A Herrmann, M Drab, and M C Cardoso. Cell Entry of Arginine-rich Peptides Is Independent of Endocytosis. *Journal of Biological Chemistry*, 284(6):3370–3378, December 2008.

[125] Falk Duchardt, Mariola Fotin-Mleczek, Heinz Schwarz, Rainer Fischer, and Roland Brock. A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic*, 8(7):848–866, July 2007.

[126] C Palm-Apergi, A Lorents, K Padari, M Pooga, and M Hallbrink. The membrane repair response masks membrane disturbances caused by cell-penetrating peptide uptake. *The FASEB Journal*, 23(1):214–223, 2009.

[127] Jean Philippe Richard, Kamran Melikov, Eric Vives, Corinne Ramos, Birgit Verbeure, Mike J Gait, Leonid V Chernomordik, and Bernard Lebleu. Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. *J Biol Chem*, 278(1): 585–590, January 2003.

[128] James R Maiolo, Elizabeth A Ottinger, and Marc Ferrer. Specific redistribution of cell-penetrating peptides from endosomes to the cytoplasm and nucleus upon laser illumination. *J Am Chem Soc*, 126(47):15376–15377, December 2004.

[129] T L Blundell, J E Pitts, I J Tickle, S P Wood, and C W Wu. X-ray analysis (1. 4-A resolution) of avian pancreatic polypeptide: Small globular protein hormone. *Proc Natl Acad Sci USA*, 78(7):4175–4179, July 1981.

[130] Abby M Hodges and Alanna Schepartz. Engineering a Monomeric Miniature Protein. *J Am Chem Soc*, 129(36):11024–11025, September 2007.

[131] J W Chin and A Schepartz. Concerted evolution of structure and function in a miniature protein. *J Am Chem Soc*, 123(12):2929–2930, March 2001.

[132] Stacey E Rutledge, Heather M Volkman, and Alanna Schepartz. Molecular Recognition of Protein Surfaces: High Affinity Ligands for the CBP KIX Domain. *J Am Chem Soc*, 125(47):14336–14347, November 2003.

[133] Daniel C Smith, J Michael Lord, Lynne M Roberts, Eric Tartour, and Ludger Johannes. 1st class ticket to class I: protein toxins as pathfinders for antigen presentation. *Traffic*, 3(10):697–704, October 2002.

[134] Douglas S Daniels and Alanna Schepartz. Intrinsically cell-permeable miniature proteins based on a minimal cationic PPII motif. *J Am Chem Soc*, 129(47):14578–14579, November 2007.

[135] D A Keire, M Kobayashi, T E Solomon, and J R Reeve. Solution structure of monomeric peptide YY supports the functional significance of the PP-fold. *Biochemistry*, 39(32):9935–9942, August 2000.

[136] Christine K Payne, Sara A Jones, Chen Chen, and Xiaowei Zhuang. Internalization and Trafficking of Cell Surface Proteoglycans and Proteoglycan-Binding Ligands. *Traffic*, 8(4):389–401, April 2007.

[137] C R Hopkins and I S Trowbridge. Internalization and processing of transferrin and the transferrin receptor in human carcinoma A431 cells. *J Cell Biol*, 97(2):508–521, August 1983.

[138] M Tyagi, M Rusnati, M Presta, and M Giacca. Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans. *J Biol Chem*, 276(5):3254–3261, February 2001.

[139] J A Hanover, M C Willingham, and I Pastan. Kinetics of transit of transferrin and epidermal growth factor through clathrin-coated membranes. *Cell*, 39(2 Pt 1): 283–293, December 1984.

[140] Andreas Vonderheit and Ari Helenius. Rab7 associates with early endosomes to mediate sorting and transport of Semliki forest virus to late endosomes. *Plos Biol*, 3(7):e233, July 2005.

[141] Betsy A Smith, Douglas S Daniels, Abigail E Coplin, Gregory E Jordan, Lynn M McGregor, and Alanna Schepartz. Minimally cationic cell-permeable miniature proteins via alpha-helical arginine display. *J Am Chem Soc*, 130(10):2948–2949, March 2008.

[142] Catherine L Watkins, Dirk Schmaljohann, Shiroh Futaki, and Arwyn T Jones. Low concentration thresholds of plasma membranes for rapid energy-independent translocation of a cell-penetrating peptide. *Biochem J*, 420(2):179–189, June 2009.

[143] Peng Yu, Bo Liu, and Thomas Kodadek. A high-throughput assay for assessing the cell permeability of combinatorial libraries. *Nat Biotechnol*, 23(6):746–751, June 2005.

[144] K L Carey, S A Richards, K M Lounsbury, and I G Macara. Evidence using a green fluorescent protein-glucocorticoid receptor chimera that the Ran/TC4 GTPase mediates an essential function independent of nuclear protein import. *J Cell Biol*, 133 (5):985–996, June 1996.

[145] Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat, Polina Golland, and David M Sabatini. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol*, 7(10):R100, 2006.

[146] Alessandra Giodini and Peter Cresswell. Hsp90-mediated cytosolic refolding of exogenous proteins internalized by dendritic cells. *EMBO J*, 27(1):201–211, January 2008.

[147] E M Abdel Ghani, S Weis, I Walev, M Kehoe, S Bhakdi, and M Palmer. Streptolysin O: inhibition of the conformational change during membrane binding of the monomer prevents oligomerization and pore formation. *Biochemistry*, 38(46): 15204–15211, November 1999.

[148] M Palmer, R Harris, C Freytag, M Kehoe, J Tranum-Jensen, and S Bhakdi. Assembly mechanism of the oligomeric streptolysin O pore: the early membrane lesion is lined by a free edge of the lipid membrane and is extended gradually during oligomerization. *EMBO J*, 17(6):1598–1605, March 1998.

[149] Iwan Walev, Michael Hombach, Wieslawa Bobkiewicz, Dominic Fenske, Sucharit Bhakdi, and Matthias Husmann. Resealing of large transmembrane pores produced by streptolysin O in nucleated cells is accompanied by NF-kappaB activation and downstream events. *The FASEB Journal*, 16(2):237–239, February 2002.

[150] I Walev, S C Bhakdi, F Hofmann, N Djonder, A Valeva, K Aktories, and S Bhakdi. Delivery of proteins into living cells by reversible membrane permeabilization with streptolysin-O. *Proc Natl Acad Sci USA*, 98(6):3185–3190, March 2001.

[151] K Y Larbi and B D Gomperts. Practical considerations regarding the use of streptolysin-O as a permeabilising agent for cells in the investigation of exocytosis. *Biosci. Rep.*, 16(1):11–21, February 1996.

[152] M J Androlewicz, K S Anderson, and P Cresswell. Evidence that transporters associated with antigen processing translocate a major histocompatibility complex class I-binding peptide into the endoplasmic reticulum in an ATP-dependent manner. *Proc Natl Acad Sci USA*, 90(19):9130–9134, October 1993.

[153] M J Androlewicz, B Ortmann, P M van Endert, T. Spies, and P Cresswell. Characteristics of peptide and major histocompatibility complex class I/beta 2-microglobulin binding to the transporters associated with antigen processing (TAP1 and TAP2). *Proc Natl Acad Sci USA*, 91(26):12716–12720, December 1994.

[154] M J Androlewicz and P Cresswell. Human transporters associated with antigen processing possess a promiscuous peptide-binding site. *Immunity*, 1(1):7–14, April 1994.

[155] M A Kehoe, L Miller, J A Walker, and G J Boulnois. Nucleotide sequence of the streptolysin O (SLO) gene: structural homologies between SLO and other membrane-damaging, thiol-activated toxins. *Infect Immun*, 55(12):3228–3232, December 1987.

[156] Chie Hotta, Haruka Fujimaki, Masahiro Yoshinari, Masatoshi Nakazawa, and Mutsuhiko Minami. The delivery of an antigen from the endocytic compartment into the cytosol for cross-presentation is restricted to early immature dendritic cells. *Immunology*, 117(1):97–107, 2006.

[157] M E Noelken, P J Chang, and J R Kimmel. Conformation and association of pancreatic polypeptide from three species. *Biochemistry*, 19(9):1838–1843, April 1980.

[158] P J Chang, M E Noelken, and J R Kimmel. Reversible dimerization of avian pancreatic polypeptide. *Biochemistry*, 19(9):1844–1849, April 1980.

[159] Andrea Balbo, Patrick H Brown, Emory H Braswell, and Peter Schuck. Measuring protein-protein interactions by equilibrium sedimentation. *Curr Protoc Immunol*, Chapter 18:Unit 18.8, November 2007.

[160] Peter Schuck. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem*, 320(1):104–124, September 2003.

[161] Jennifer Vistica, Julie Dam, Andrea Balbo, Emine Yikilmaz, Roy A Mariuzza, Tracey A Rouault, and Peter Schuck. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem*, 326(2):234–256, March 2004.

[162] James L Cole. Analysis of heterogeneous interactions. *Methods in Enzymology*, 384: 212–232, 2004.

[163] H Durchschlag and P Zipper. Calculation of the partial volume of organic compounds and polymers. *Ultracentrifugation*, 94:20–39, 1994.

[164] Thomas Wollert, Christian Wunder, Jennifer Lippincott-Schwartz, and James H Hurley. Membrane scission by the ESCRT-III complex. *Nature*, 457(7235):172–177, March 2009.

[165] Tadeusz Muzioł, Estela Pineda-Molina, Raimond B Ravelli, Alessia Zamborlini, Yoshiko Usami, Heinrich Göttlinger, and Winfried Weissenhorn. Structural basis for budding by the ESCRT-III factor CHMP3. *Developmental Cell*, 10(6):821–830, June 2006.

[166] Steven Johnson, Pietro Roversi, Marianela Espina, Andrew Olive, Janet E Deane, Susan Birket, Terry Field, William D Picking, Ariel J Blocker, Edouard E Galyov, Wendy L Picking, and Susan M Lea. Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD. *J Biochem*, 282(6):4035–4044, February 2007.

[167] Constantina Bakolitsa, Daniel M Cohen, Laurie A Bankston, Andrey A Bobkov, Gregory W Cadwell, Lisa Jennings, David R Critchley, Susan W Craig, and Robert C Liddington. Structural basis for vinculin activation at sites of cell adhesion. *Nature*, 430(6999):583–586, July 2004.

[168] Pietro De Camilli, Hong Chen, Joel Hyman, Ezequiel Panepucci, Alex Bateman, and Axel T Brunger. The ENTH domain. *FEBS Lett*, 513(1):11–18, February 2002.

[169] Toshiki Itoh and Pietro De Camilli. BAR, F-BAR (EFC) and ENTH/ANTH domains in the regulation of membrane-cytosol interfaces and membrane curvature. *Biochim Biophys Acta*, 1761(8):897–912, August 2006.

[170] Robert V Stahelin, Fei Long, Brian J Peter, Diana Murray, Pietro De Camilli, Harvey T McMahon, and Wonhwa Cho. Contrasting membrane interaction mechanisms of AP180 N-terminal homology (ANTH) and epsin N-terminal homology (ENTH) domains. *J Biochem*, 278(31):28993–28999, August 2003.

[171] Yong-Uk Kwon and Thomas Kodadek. Quantitative Evaluation of the Relative Cell Permeability of Peptoids and Peptides. *J Am Chem Soc*, 129(6):1508–1509, February 2007.

[172] Raheem Peerani, Balaji M Rao, Celine Bauwens, Ting Yin, Geoffrey A Wood, Andras Nagy, Eugenia Kumacheva, and Peter W Zandstra. Niche-mediated control of human embryonic stem cell self-renewal and differentiation. *EMBO J*, 26(22): 4744–4755, October 2007.

[173] G J Ding, P A Fischer, R C Boltz, J A Schmidt, J J Colaianne, A Gough, R A Rubin, and D K Miller. Characterization and quantitation of NF-kappaB nuclear translocation induced by interleukin-1 and tumor necrosis factor-alpha. Development and use of a high capacity fluorescence cytometric system. *J Biol Chem*, 273(44):28897–28905, October 1998.

[174] S C Gill and P H von Hippel. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, 182(2):319–326, November 1989.

[175] Yuki Takechi, Haruka Yoshii, Masafumi Tanaka, Toru Kawakami, Saburo Aimoto, and Hiroyuki Saito. Physicochemical mechanism for the enhanced ability of lipid membrane penetration of polyarginine. *Langmuir*, 27(11):7099–7107, June 2011.

[176] Fatouma Said Hassane, Rachida Abes, Samir El Andaloussi, Taavi Lehto, Rannar Sillard, Ulo Langel, and Bernard Lebleu. Insights into the cellular trafficking of splice redirecting oligonucleotides complexed with chemically modified cell-penetrating peptides. *J Control Release*, 153(2):163–172, July 2011.

[177] André Ziegler, Pierluigi Nervi, Markus Dürrenberger, and Joachim Seelig. The cationic cell-penetrating peptide CPP(TAT) derived from the HIV-1 protein TAT is rapidly transported into living fibroblasts: optical, biophysical, and metabolic evidence. *Biochemistry*, 44(1):138–148, January 2005.

[178] Henry D Herce and Angel E Garcia. Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes. *Proceedings of the National Academy of Sciences*, 104(52):20805–20810, December 2007.

[179] H D Herce, A E Garcia, J Litt, R S Kane, P Martin, N Enrique, A Rebolledo, and V Milesi. Arginine-rich peptides destabilize the plasma membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides. *Biophys J*, 97(7):1917–1925, October 2009.

[180] Guillaume Drin, Sylvine Cottin, Emmanuelle Blanc, Anthony R Rees, and Jamal Temsamani. Studies on the internalization mechanism of cationic cell-penetrating peptides. *J Biol Chem*, 278(33):31192–31201, August 2003.

[181] I Mellman. Endocytosis and molecular sorting. *Annu Rev Cell Dev Biol*, 12:575–625, 1996.

[182] Satyajit Mayor and Richard E Pagano. Pathways of clathrin-independent endocytosis. *Nat Rev Mol Cell Biol*, 8(8):603–612, August 2007.

[183] Sean D Conner and Sandra L Schmid. Regulated portals of entry into the cell. *Nature*, 422(6927):37–44, March 2003.

[184] Julie A Champion, Amanda Walker, and Samir Mitragotri. Role of particle size in phagocytosis of polymeric microspheres. *Pharm Res*, 25(8):1815–1821, August 2008.

[185] Shirley M L Tse, Wendy Furuya, Elizabeth Gold, Alan D Schreiber, Kirsten Sandvig, Robert D Inman, and Sergio Grinstein. Differential role of actin, clathrin, and dynamin in Fc gamma receptor-mediated endocytosis and phagocytosis. *J Biochem*, 278(5):3331–3338, January 2003.

[186] Jason M Kinchen, Kimon Doukoumetzidis, Johann Almendinger, Lilli Stergiou, Annie Tosello-Trampont, Costi D Sifri, Michael O Hengartner, and Kodi S Ravichandran. A pathway for phagosome maturation during engulfment of apoptotic cells. *Nat Cell Biol*, 10(5):556–566, April 2008.

[187] Kassidy K Huynh and Sergio Grinstein. Phagocytosis: dynamin's dual role in phagosome biogenesis. *Curr. Biol.*, 18(13):R563–5, July 2008.

[188] J Hirst and M S Robinson. Clathrin and adaptors. *Biochim Biophys Acta*, 1404(1-2): 173–193, August 1998.

[189] M S Robinson. The role of clathrin, adaptors and dynamin in endocytosis. *Curr Opin Cell Biol*, 6(4):538–544, August 1994.

[190] C Harding, J Heuser, and P Stahl. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. *J Cell Biol*, 97(2):329–339, August 1983.

[191] Lucas Pelkmans, Daniel Püntener, and Ari Helenius. Local actin polymerization and dynamin recruitment in SV40-induced internalization of caveolae. *Science*, 296 (5567):535–539, April 2002.

[192] Lucas Pelkmans and Ari Helenius. Endocytosis via caveolae. *Traffic*, 3(5):311–320, May 2002.

[193] Lucas Pelkmans, Thomas Bürli, Marino Zerial, and Ari Helenius. Caveolin-stabilized membrane domains as multifunctional transport and sorting devices in endocytic membrane traffic. *Cell*, 118(6):767–780, September 2004.

[194] Peter Thomsen, Kirstine Roepstorff, Martin Stahlhut, and Bo van Deurs. Caveolae are highly immobile plasma membrane microdomains, which are not involved in constitutive endocytic trafficking. *Molecular Biology of the Cell*, 13(1):238–250, 2002.

[195] Arnold Hayer, Miriam Stoeber, Danilo Ritz, Sabrina Engel, Hemmo H Meyer, and Ari Helenius. Caveolin-1 is ubiquitinated and targeted to intralumenal vesicles in endolysosomes for degradation. *J Cell Biol*, 191(3):615–629, November 2010.

[196] S Engel, T Heger, R Mancini, F Herzog, J Kartenbeck, A Hayer, and A Helenius. Role of Endosomes in Simian Virus 40 Entry and Infection. *J Virol*, 85(9):4198–4211, April 2011.

[197] R Schekman and L Orci. Coat proteins and vesicle budding. *Science*, 271(5255): 1526–1533, March 1996.

[198] Shawn M Ferguson, Gabor Brasnjo, Mitsuko Hayashi, Markus Wölfel, Chiara Collesi, Silvia Giovedi, Andrea Raimondi, Liang-Wei Gong, Pablo Ariel, Summer Paradise, Eileen O'toole, Richard Flavell, Ottavio Cremona, Gero Miesenböck, Timothy A Ryan, and Pietro De Camilli. A selective activity-dependent requirement for dynamin 1 in synaptic vesicle endocytosis. *Science*, 316(5824):570–574, April 2007.

[199] Shawn Ferguson, Andrea Raimondi, Summer Paradise, Hongying Shen, Kumi Mesaki, Agnes Ferguson, Olivier Destaing, Genevieve Ko, Junko Takasaki, Ottavio Cremona, Eileen O Toole, and Pietro De Camilli. Coordinated Actions of

Actin and BAR Proteins Upstream of Dynamin at Endocytic Clathrin-Coated Pits. *Developmental Cell*, 17(6):811–822, December 2009.

[200] Zhi-jie Cheng, Raman Deep Singh, Deepak K Sharma, Eileen L Holicky, Kentaro Hanada, David L Marks, and Richard E Pagano. Distinct mechanisms of clathrin-independent endocytosis have unique sphingolipid requirements. *Molecular Biology of the Cell*, 17(7):3197–3210, July 2006.

[201] Naava Naslavsky, Roberto Weigert, and Julie G Donaldson. Characterization of a nonclathrin endocytic pathway: membrane cargo and lipid requirements. *Molecular Biology of the Cell*, 15(8):3542–3552, August 2004.

[202] Rahul Chadda, Mark T Howes, Sarah J Plowman, John F Hancock, Robert G Parton, and Satyajit Mayor. Cholesterol-sensitive Cdc42 activation regulates actin polymerization for endocytosis via the GEEC pathway. *Traffic*, 8(6):702–717, June 2007.

[203] A Subtil, I Gaidarov, K Kobylarz, M A Lampson, J H Keen, and T E McGraw. Acute cholesterol depletion inhibits clathrin-coated pit budding. *Proc Natl Acad Sci USA*, 96(12):6775–6780, June 1999.

[204] S K Rodal, G Skretting, O Garred, F Vilhardt, B van Deurs, and K Sandvig. Extraction of cholesterol with methyl-beta-cyclodextrin perturbs formation of clathrin-coated endocytic vesicles. *Molecular Biology of the Cell*, 10(4):961–974, April 1999.

[205] Roberto Zoncu, Rushika M Perera, Daniel M Balkin, Michelle Pirruccello, Derek Toomre, and Pietro De Camilli. A phosphoinositide switch controls the maturation and signaling properties of APPL endosomes. *Cell*, 136(6):1110–1121, March 2009.

[206] J P Gorvel, P Chavrier, M Zerial, and J Gruenberg. rab5 controls early endosome fusion in vitro. *Cell*, 64(5):915–925, March 1991.

[207] C Bucci, R G Parton, I H Mather, H Stunnenberg, K Simons, B Hoflack, and M Zerial. The small GTPase rab5 functions as a regulatory factor in the early endocytic pathway. *Cell*, 70(5):715–728, September 1992.

[208] Nils C Gauthier, Pascale Monzo, Vincent Kaddai, Anne Doye, Vittorio Ricci, and Patrice Boquet. Helicobacter pylori VacA cytotoxin: a probe for a clathrin-independent and Cdc42-dependent pinocytic pathway routed to late endosomes. *Molecular Biology of the Cell*, 16(10):4852–4866, October 2005.

[209] Naava Naslavsky, Roberto Weigert, and Julie G Donaldson. Convergence of non-clathrin- and clathrin-derived endosomes involves Arf6 inactivation and changes in phosphoinositides. *Molecular Biology of the Cell*, 14(2):417–431, February 2003.

[210] Manjula Kalia, Sudha Kumari, Rahul Chadda, Michelle M Hill, Robert G Parton, and Satyajit Mayor. Arf6-independent GPI-anchored protein-enriched early endosomal compartments fuse with sorting endosomes via a Rab5/phosphatidylinositol-3'-kinase-dependent machinery. *Molecular Biology of the Cell*, 17(8):3689–3704, August 2006.

[211] Manojkumar A Puthenveedu, Benjamin Lauffer, Paul Temkin, Rachel Vistein, Peter Carlton, Kurt Thorn, Jack Taunton, Orion D Weiner, Robert G Parton, and Mark von Zastrow. Sequence-Dependent Sorting of Recycling Proteins by Actin-Stabilized Endosomal Microdomains. *Cell*, 143(5):761–773, November 2010.

[212] Jochen Rink, Eric Ghigo, Yannis Kalaidzidis, and Marino Zerial. Rab conversion as a mechanism of progression from early to late endosomes. *Cell*, 122(5):735–749, September 2005.

[213] Dmitry Poteryaev, Sunando Datta, Karin Ackema, Marino Zerial, and Anne Spang. Identification of the Switch in Early-to-Late Endosome Transition. *Cell*, 141(3):497–508, January 2010.

[214] Jatta Huotari and Ari Helenius. Endosome maturation. *EMBO J*, 30(17):3481–3500, August 2011.

[215] M Lakadamyali, M Rust, and X Zhuang. Ligands for Clathrin-Mediated Endocytosis Are Differentially Sorted into Distinct Populations of Early Endosomes. *Cell*, 124(5):997–1009, March 2006.

[216] Eric Macia, Marcelo Ehrlich, Ramiro Massol, Emmanuel Boucrot, Christian Brunner, and Tomas Kirchhausen. Dynasore, a Cell-Permeable Inhibitor of Dynamin. *Developmental Cell*, 10(6):839–850, June 2006.

[217] Gary J Doherty and Harvey T McMahon. Mechanisms of Endocytosis. *Annu. Rev. Biochem.*, 78(1):857–902, June 2009.

[218] Hong Cao, Jing Chen, Muyiwa Awoniyi, John R Henley, and Mark A McNiven. Dynamin 2 mediates fluid-phase micropinocytosis in epithelial cells. *J Cell Sci*, 120 (Pt 23):4167–4177, December 2007.

[219] Ya-Wen Liu, Mark C Surka, Thomas Schroeter, Vasyl Lukiyanchuk, and Sandra L Schmid. Isoform and splice-variant specific functions of dynamin-2 revealed by analysis of conditional knock-out cells. *Molecular Biology of the Cell*, 19(12):5347–5359, December 2008.

[220] Mirkka Koivusalo, Christopher Welch, Hisayoshi Hayashi, Cameron C Scott, Moshe Kim, Todd Alexander, Nicolas Touret, Klaus M Hahn, and Sergio Grinstein. Amiloride inhibits macropinocytosis by lowering submembranous pH and preventing Rac1 and Cdc42 signaling. *J Cell Biol*, 188(4):547–563, February 2010.

[221] Maria D Moya-Ortega, Carmen Alvarez-Lorenzo, Hákon H Sigurdsson, Angel Concheiro, and Thorsteinn Loftsson. Gamma-Cyclodextrin hydrogels and semi-interpenetrating networks for sustained delivery of dexamethasone. *Carbohydrate Polymers*, 80(3):900–907, May 2010.

[222] T Yoshimori, A Yamamoto, Y Moriyama, M Futai, and Y Tashiro. Bafilomycin A1, a specific inhibitor of vacuolar-type H(+)-ATPase, inhibits acidification and protein degradation in lysosomes of cultured cells. *J Biochem*, 266(26):17707–17712, September 1991.

[223] E J Bowman, A Siebers, and K Altendorf. Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells. *Proc Natl Acad Sci USA*, 85(21):7972–7976, November 1988.

[224] S Christoforidis, M Miaczynska, K Ashman, M Wilm, L Zhao, S C Yip, M D Waterfield, J M Backer, and M Zerial. Phosphatidylinositol-3-OH kinases are Rab5 effectors. *Nat Cell Biol*, 1(4):249–252, August 1999.

[225] H Stenmark, R G Parton, O Steele-Mortimer, A Lütcke, J Gruenberg, and M Zerial. Inhibition of rab5 GTPase activity stimulates membrane fusion in endocytosis. *EMBO JOURNAL*, 13(6):1287–1296, March 1994.

[226] S Christoforidis, H M McBride, R D Burgoyne, and M Zerial. The Rab5 effector EEA1 is a core component of endosome docking. *Nature*, 397(6720):621–625, February 1999.

[227] G Cantalupo, P Alifano, V Roberti, C B Bruni, and C Bucci. Rab-interacting lysosomal protein (RILP): the Rab7 effector required for transport to lysosomes. *EMBO JOURNAL*, 20(4):683–693, February 2001.

[228] Qiming Sun, Wiebke Westphal, Kwun Ngok Wong, Irena Tan, and Qing Zhong. Rubicon controls endosome maturation as a Rab7 effector. *Proc Natl Acad Sci USA*, 107(45):19338–19343, November 2010.

[229] W Querbes, B A O'Hara, G Williams, and W J Atwood. Invasion of host cells by JC virus identifies a novel role for caveolae in endosomal sorting of noncaveolar ligands. *J Virol*, 80(19):9402–9413, October 2006.

[230] N B Cole, C L Smith, N Sciaky, M Terasaki, M Edidin, and J Lippincott-Schwartz. Diffusional mobility of Golgi proteins in membranes of living cells. *Science*, 273 (5276):797–801, August 1996.

[231] G Patterson, K Hirschberg, R Polishchuk, D Gerlich, R Phair, and J Lippincottschwartz. Transport through the Golgi Apparatus by Rapid Partitioning within a Two-Phase Membrane System. *Cell*, 133(6):1055–1067, June 2008.

[232] Harpreet Kaur, Aarti Garg, and G P S Raghava. PEPstr: a de novo method for tertiary structure prediction of small bioactive peptides. *Protein Pept. Lett.*, 14(7): 626–631, 2007.

[233] Christophe Combet, Martin Jambon, Gilbert Deléage, and Christophe Geourjon. Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics*, 18 (1):213–214, January 2002.

[234] Barbara M Tynan-Connolly and Jens Erik Nielsen. pKD: re-designing protein pKa values. *Nucleic Acids Research*, 34(Web Server issue):W48–51, July 2006.

[235] Gaël Vidricaire and Michel J Tremblay. Rab5 and Rab7, but not ARF6, govern the early events of HIV-1 infection in polarized human placental cells. *J Immunol*, 175 (10):6517–6530, November 2005.

[236] A Puhar, E A Johnson, O Rossetto, and C Montecucco. Comparison of the pH-induced conformational change of different clostridial neurotoxins. *Biochem Biophys Res Commun*, 319(1):66–71, June 2004.

[237] JP Shaffer. Multiple hypothesis testing. *Annual Review of Psychology*, 46:561–584, February 1995.

[238] Carolyn B Coyne, Le Shen, Jerrold R Turner, and Jeffrey M Bergelson. Coxsackievirus entry across epithelial tight junctions requires occludin and the small GTPases Rab34 and Rab5. *Cell Host Microbe*, 2(3):181–192, September 2007.

[239] Priti Kumar, Haoquan Wu, Jodi L Mcbride, Kyeong-Eun Jung, Moon Hee Kim, Beverly L Davidson, Sang Kyung Lee, Premlata Shankar, and N Manjunath. Transvascular delivery of small interfering RNA to the central nervous system. *Nature*, 448 (7149):39–43, July 2007.