ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

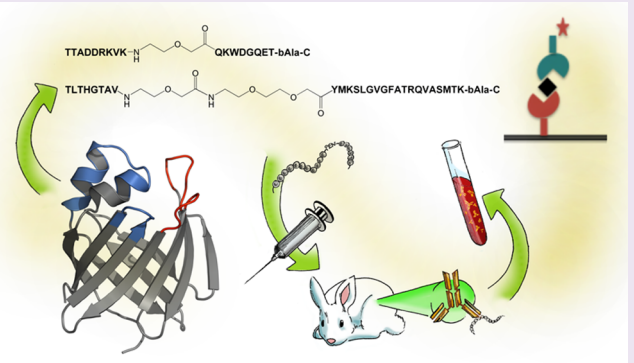
Rational Epitope Design for Protein Targeting

Claudio Peri, Paola Gagni, Fabio Combi, Alessandro Gori, Marcella Chiari, Renato Longhi, Marina Cretich, and Giorgio Colombo*

Isituto di Chimica del Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano, Italy

Supporting Information

ABSTRACT: We present a new multidisciplinary strategy integrating computational biology with high-throughput microarray analysis aimed to translate molecular understanding of protein-antibody recognition into the design of efficient and selective protein-based analytical and diagnostic tools. The structures of two proteins with different folds and secondary structure contents, namely, the beta-barrel FABP and the α-helical S100B, were used as the basis for the prediction and design of potential antibody-binding epitopes using the recently developed MLCE computational method. Starting from the idea that the structure, dynamics, and stability of a protein-antigen play a key role in the interaction with antibodies, MLCE integrates the analysis of the dynamical and energetic properties of proteins to identify nonoptimized, low-intensity energetic interaction-networks on the surface of the isolated antigens, which correspond to substructures that can aptly be recognized by a binding partner. The identified epitopes were next synthesized as free peptides and used to elicit specific antibodies in rabbits. Importantly, the resulting antibodies were proven to specifically and selectively recognize the original, full-length proteins in microarray-based tests. Competition experiments further demonstrated the specificity of the molecular recognition between the target immobilized proteins and the generated antibodies. Our integrated computational and microarray-based results demonstrate the possibility to rationally discover and design synthetic epitopes able to elicit antibodies specific for full-length proteins starting only from three-dimensional structural information on the target. We discuss implications for diagnosis and vaccine development purposes.



The identification of protein parts that are crucial for protein−protein and protein−inhibitor interactions is essential to understand biological systems and can help illuminate useful targets for pharmacological and diagnostic intervention. Interference with the exposition of protein segments engaged in disease-related pathways could indeed reveal the molecular and structural determinants of a pathological condition and, on this basis, facilitate the development of strategies apt to blocking the permissibility of pathogenic interactions. In this context, the engineering of novel proteins and peptides that display recognition properties similar to the ones of natural protein binding sites is a long-standing goal in biological chemistry. This capability is particularly relevant in the fields of diagnostics and vaccine development, especially for diseases for which no effective marker or treatment exist. In both fields, the use of designed antigenic molecules for the elicitation of antibodies as biomarkers and inhibitors of disease states is an important issue. Antibodies indeed function by targeting molecules that are characteristic of a certain condition. Antibodies of great utility have been identified[1] with a deep impact on the diagnosis and treatment of autoimmune disorders,[2] infections,[3] and cancers.[4] They function by binding to specific antigen substructures, called epitopes. The increasing knowledge generated by reverse vaccinology and structural genomics programs has revealed that the structure, dynamics, and stability of a protein domain play a key role in determining the antigenic properties of specific (fragments of) protein constructs, in the interaction of epitopes with antibodies, and finally in their relevance for a protective response.[5−8] The identification of reactive antigen parts determinant for recognition and binding may thus facilitate the generation of useful antibodies. Such considerations suggest new opportunities for the development of rational approaches aimed at discovering and optimizing epitope sequences from protein antigens. In this framework, the combination of experimental atomic-resolution information with computational biology may open new possibilities to design and manipulate molecules able to elicit and/or bind specific antibodies against a certain target. Successful realizations of these principles have been illustrated by Schief and co-workers,[9,10] who demonstrated the elicitation of neutralizing antibodies against the HIV-1 gp41 epitope based on the use of computational biology techniques to guide the grafting of epitopes onto selected acceptor scaffolds. In this

Received: September 12, 2012
Accepted: November 8, 2012
Published: November 8, 2012

dx.doi.org/10.1021/cb300487u | ACS Chem. Biol. 2013, 8, 397−404

work, we present a new multidisciplinary strategy integrating computational biology with high-throughput microarray analysis to translate the molecular understanding of the determinants of protein-antibody recognition and the prediction of antibody-binding regions of proteins into the design of new efficient protein-based diagnostic tools to reveal specific biomarkers and to select-validate antigens for vaccine development. Starting from the structures of two different test proteins that have been proposed as markers for neurological disorders and damage, namely, the beta-barrel FABP and the α-helical S100B[11,12] (Figure 1), a recently developed computational method is used to identify and guide the design of potential epitopes.[13,14] The designed epitopes, synthesized as free peptides and coupled to the carrier protein KLH (Keyhole Limpet Hemocyanin) are then used to elicit specific antibodies in rabbits. The resulting antibodies are then proved to specifically recognize the original, full-length proteins in microarray-based tests. The results of this endeavor may have an impact at the level of basic and applied molecular medicine, providing a new protocol in which the atomic details of antigen molecular recognition properties are elucidated and exploited to design new experimentally useful protein-, peptide-, or antibody-based microarrays. The method presented here is of general application and can be extended to other protein antigens for diagnosis and vaccine development purposes.

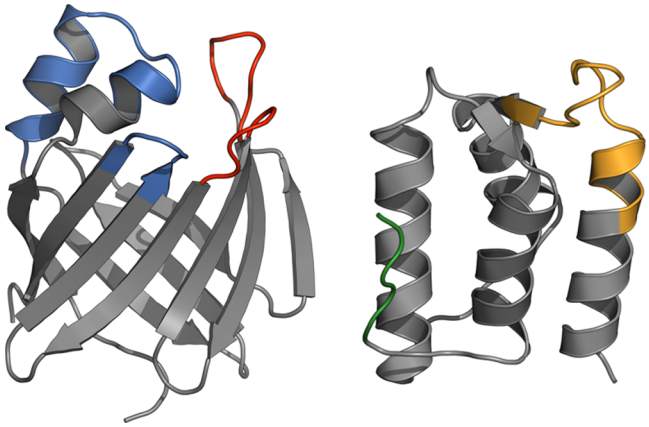


Figure 1. Epitope mapping on FABP and S100B structures. Left panel: secondary structure representation of FABP. The two loops comprising the predicted epitope **f-ep1** are highlighted in red. Epitope **f-ep2**, comprising the helical portion of the protein and the nearby β-turn, is colored blue. Right panel: cartoon representation of S100B, secondary structure. The predicted epitope **s-ep1** is colored yellow. Epitope **s-ep2** is colored green.

RESULTS

To select and guide the synthesis of optimal peptide sequences able to recapitulate the antibody binding properties of FABP and S100B, we applied the recently developed MLCE (matrix of local coupling energies) method.[13] This computational method has been developed and validated to predict antibody-binding sites on protein antigens from the analysis of molecular dynamics (MD) simulations. The idea at the basis of the approach is that epitopes correspond to localized regions with nonoptimized energetic couplings with the rest of the protein allowing them to undergo conformational changes, to be recognized by a binding partner and to tolerate mutations with minimal energetic expense. Indeed, protein antigens should be able to tolerate mutations that help the pathogen evade the immune defense system of the host without impairing the stability of the protein, which is required for its functionality. The MLCE method proved able to efficiently predict discontinuous as well as linear epitopes based only on the structural and physicochemical properties of the antigen,[13,14] without requiring any previous knowledge on antibody binding of related homologues, or training with a data set of known sequences, geometric descriptors, antibody−protein interactions, etc. The detailed results of the simulations, the results of the cluster analysis, the representative structures of the main clusters, and the results of MLCE analysis are presented as Supporting Information.

Epitope Prediction, Design, and Synthesis of Antigenic Peptides. The MLCE epitope prediction method was applied to representative structures obtained from MD simulations of FABP and S100B (see Materials and Methods). The results are presented in Figure 1. Two epitopes on each protein were predicted.

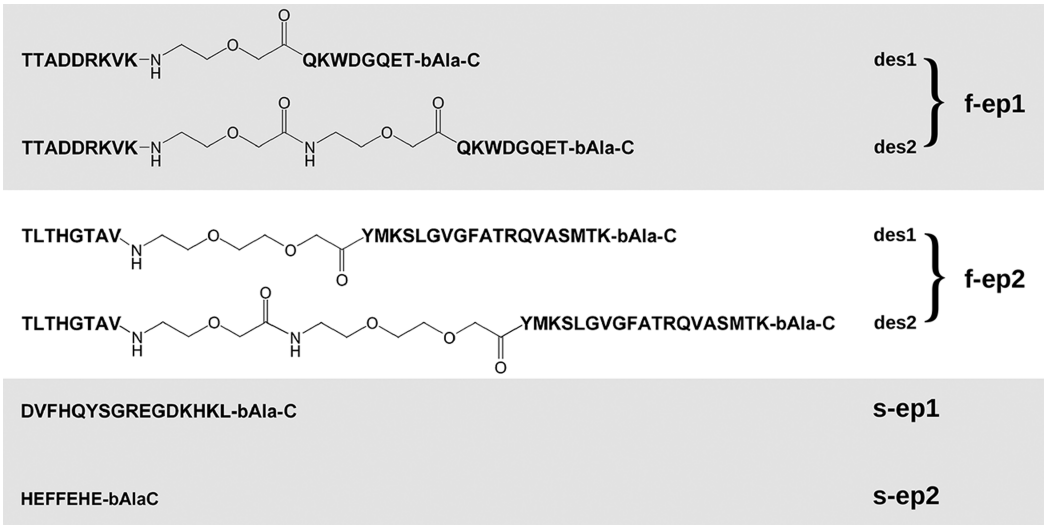


Figure 2. Synthetic epitopes. Sequences and structures of the synthesized epitopes, designed on the basis of computational epitope prediction.

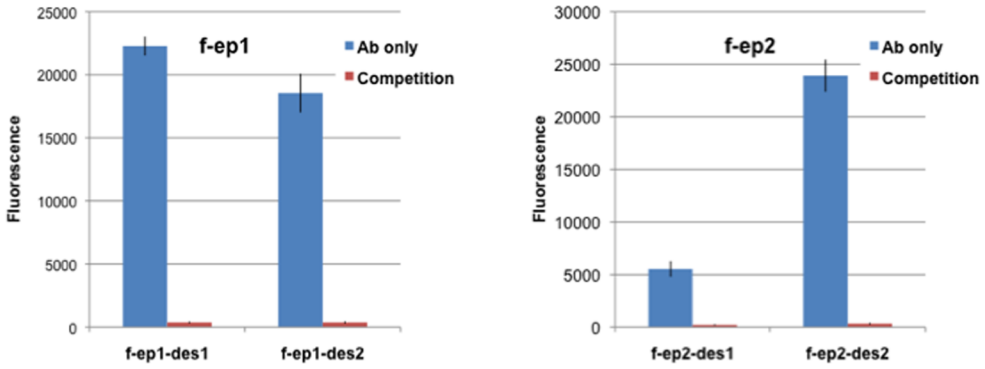


Figure 3. Microarray-based tests for FABP. Average fluorescence intensity of human FABP spots upon incubation with Cy3-labeled antibodies elicited by **f-ep1** (left panel) and **f-ep2** (right panel). Blue bars represent interaction between immobilized FABP and antibodies elicited by the peptide variants. Specificity of interaction is demonstrated by the quenching of the fluorescence signal observed when antibodies are challenged with the eliciting peptides (red bars).

In the case of FABP, the prediction returned two conformational epitopes. The first one, labeled **f-ep1** is composed of two loops (with sequences T**TADDRK**VK and QK**WDGQE**T, respectively) connecting beta-sheets res. 71−74 and res. 80−88; the second one, labeled **f-ep2**, spans the exposed helical part of the protein (sequence Y**MKSLGVG-FATRQVASMTK**) plus the short loop (sequence TL**THGTA**V) connecting and including partial beta-sheets res. 113−120 and res. 123−131. The two predicted epitopes on FABP are located in the same region of the protein (see Figure 1).

Next, we set out to design peptide-based mimics able to display the constitutive parts of each of the two conformational epitopes to be used in the experimental elicitations of antibodies. To this end, the results of MD simulations of FABP were analyzed to select the optimal linkers for the connection of the chains forming either **f-ep1** or **f-ep2**. The average value and time-evolution of the distance between the terminals of the sequences that define each conformational epitope were analyzed. To connect the terminals, keeping the sequences in an orientation akin to the one they featured in the full-length FABP protein, we chose water-soluble, nonimmunogenic linkers that could also increase the bioavailability of the resulting molecules, an important requirement in the development of efficient biotherapeutics. In the case of **f-ep1**, the average distance between K82 at the N-terminal and Q96 at the C-terminal was 5.15 Å. A single **Pen** (PEN: 3-oxa-pentanoic acid, Figure 2) moiety was found to be sufficient to mimic the average distance and ensure the correct orientation of the two sequence stretches. Therefore, the results of the structure-based prediction were translated into the synthesis of the first designed epitope with sequence: T**TADDRK**VK-**Pen**-QK**WDGQE**T-bAlaC, labeled **f-ep1-des1** (Figure 2). A more flexible variant of the peptide was also prepared with the inclusion of two **Pen** moieties resulting in the following sequence T**TADDRK**VK-**PenPen**-QK**WDGQE**T-bAlaC, labeled **f-ep1-des2**. Indeed, in some cases, flexibility has been shown to correlate with immunogenicity. Moreover, in a conformational selection framework, increased flexibility may speed up the peptide's search for the structures that are required for antibody recognition and binding.[15−18] In the case of **f-ep2**, the distance between the terminals of the constituting sequences (7.48 Å) was found to be optimally approximated by an O2Oc (O2Oc: 8-amino-3,6-dioxaoctanoic acid, Figure 2) moiety. This resulted in the design of **f-ep2-des1**, TL**THGTA**V-O2Oc-Y**MKSLGVGFATRQVASMT**K-bAC (Figure 2). Also in this case, we decided to synthesize a more flexible variant with the inclusion of an additional **Pen** group, generating **f-ep2-des2**, TL**THGTA**V-**PenO2Oc**-Y**MKSLGVG-FATRQVASMT**K-bAC.

In S100B, the MLCE prediction identified as an optimal epitope candidate the linear stretch corresponding to sequence **DVFHQYSGREGDKHKL**, labeled **s-ep1.** A second epitope, corresponding to sequence **HEFFEHE**, labeled **s-ep2** (Figure 2) was predicted and synthesized. **s-ep2** spans part of α-helix-4 and the C-terminal part of the protein.

It is worth underlining that **s-ep1** and **s-ep2** are located on opposite sides of the protein, suggesting the possibility that two different antibodies developed against the two designed epitopes may bind the protein at the same time, avoiding reciprocal steric hindrance. In the light of this observation, the two antibodies can be used as sequential binders of the same substrate, leading to the formation of a ternary complex. This, in turn, may greatly increase the specificity and selectivity for the recognition of the target S100B protein.

Production of Antibodies. To evaluate the immunogenicity of the synthetic epitope peptides and their ability to elicit antibodies against their parent full-length proteins, rabbits were immunized with the synthetic peptides conjugated to KLH. Each of the immunogens induced a strong antibody response as revealed by ELISA binding tests of rabbit sera on RSA (Rabbit Serum Albumin) conjugated immunogens performed at day 42. Rabbit sera were collected and specific antibodies purified by immuno-chromatography on Sepharose conjugated to the eliciting peptides as described in the Materials and Methods section.

Microarray Tests. *Specific Detection of FABP with Designed-Peptide Elicited FABP Antibodies.* Human FABP and several nonrelated control proteins (Human Serum Albumin, Ovoalbumin, Ovomucoid, Rabbit, Mouse, and Goat Immunoglobulins) were spotted and immobilized on copoly-(DMA-NAS-MAPS) coated silicon slides.[19] The protein arrays were then incubated with Cy3-labeled-antibodies specific for **f-ep1-des1**, **f-ep1-des2**, **f-ep2-des1**, and **f-ep2-des2**. Each of the four antibodies provided a clear fluorescent signal in correspondence of FABP spots due to recognition of the full-length protein immobilized on the surface; control proteins were not recognized by the antibodies, whereas a slight nonspecific interaction was detected on immunoglobulins spots.[20] The latter is a known drawback in antibody

dx.doi.org/10.1021/cb300487u | *ACS Chem. Biol.* 2013, 8, 397−404

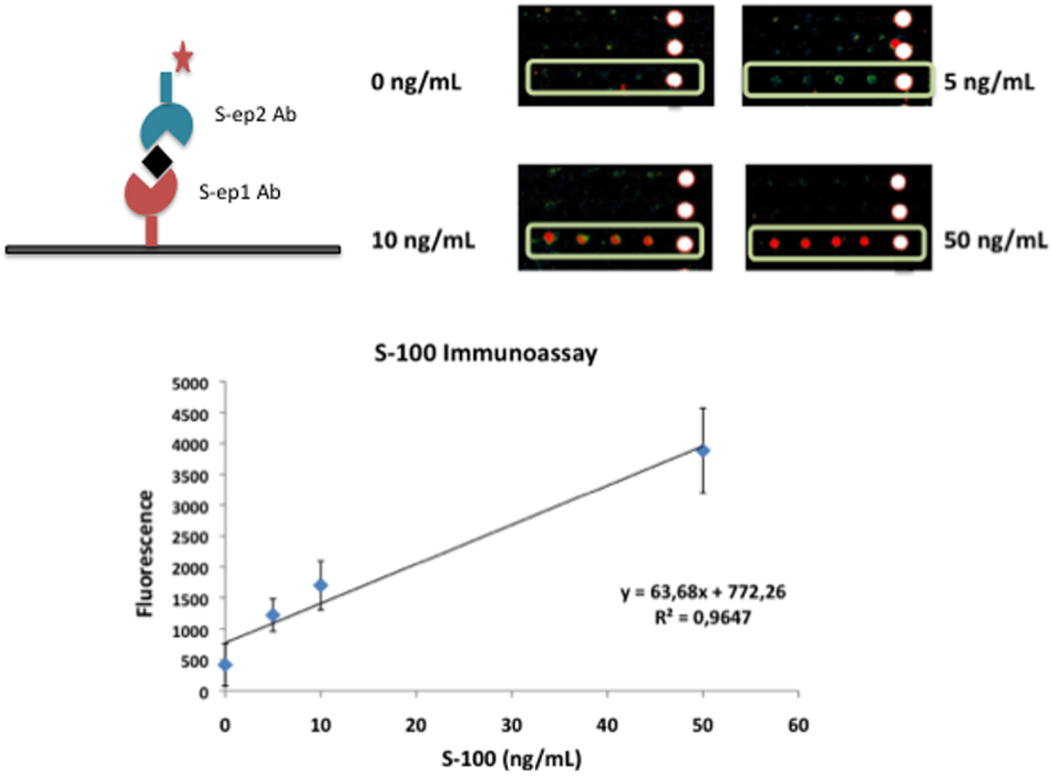


Figure 4. Microarray-based tests for S100. Upper panel: scheme of the sandwich immunoassay for S-100 detection and fluorescence images of microarrays incubated with increasing concentrations of S-100 (from 0 to 50 ng/mL). Lower panel: calibration curve for S-100 sandwich immunoassay. The extrapolated limit of detection (LOD) is 16.06 ng/mL.

microarrays. In Figure 3, the average fluorescence signal on FABP spots, upon incubation with labeled antibodies, is shown. Fluorescence generated by the antibodies directed against epitope 1 (**f-ep1**) is shown in the left panel (blue bars), molecular recognition of epitope 2 (**f-ep2**) is shown in the right panel (blue bars). For epitope 1, the use of a more flexible variant of the synthetic peptides with the introduction of a double **Pen** moiety did not provide any significant enhancement of immunogenity since fluorescence detected upon incubation with antibodies directed against **f-ep1-des2** is similar to the one provided by antibodies against **f-ep1-des1**. On the contrary, for epitope 2, the use of a more flexible peptidic variant with the introduction of the linker **PenO2OC** strongly enhanced molecular recognition of FABP as demonstrated by the higher fluorescence intensity provided by the antibodies against **f-ep2-des2**. In a competition assay, the generated fluorescent antibodies were challenged with equimolar solutions of the corresponding eliciting peptides. In every case, the fluorescence signal on FABP spots was completely quenched (Figure 3, red bars) demonstrating the specificity of the molecular recognition between the target FABP immobilized on the surface and the generated antibodies. An unrelated scrambled peptide was not effective in the competition and did not affect the protein array tests (data not shown). Worthy of note, the nonspecific interactions between the generated antibodies and the immobilized immunoglobulins were not affected by the competitive experiments still showing a constant slight fluorescent signal. Summarizing, the protein microarray tests demonstrated the ability of the designed synthetic epitopes to elicit antibodies specific for full-length human FABP.

Specific Detection of S100 with Designed-Peptide Elicited S100 Antibodies. The MLCE prediction of epitopes in S-100 identified two linear sequences localized into distinct areas of the protein. The distance between the two epitopes (on the average around 25 Å) led us to evaluate the possible use of the generated antibodies in a matched pair for a sandwich immunoassay aimed not only to detect but also to quantify S-100. The scheme of the assay is depicted in Figure 4. The antibody directed against S-100 epitope 2 (**s-ep2**) was spotted and immobilized on copoly(DMA-NAS-MAPS) coated silicon slides and used as the capture antibody; the anti-**s-ep1** antibody was labeled with Cy3 and used as the detection antibody. Figure 4 shows fluorescence of the protein array on which anti-**s-ep2** antibody was spotted together with two nonrelated rabbit antibodies. The antibody array was then incubated with several concentrations of S-100 spiked in PBS buffer from 0 to 100 ng/mL. The protein was detected using the fluorescent anti-**s-ep1** antibody. The corresponding fluorescence signals are shown in Figure 4, upper panel, highlighted rectangles. The sandwich microimmunoassay for S-100 using the generated antibodies provided a linear dose-dependent fluorescence response leading to a limit of detection for S-100 of 16.06 ng/mL as shown by the calibration curve in Figure 4, lower panel. In summary, our results demonstrate that two antibodies directed against distinct epitopes of the protein S-100 were successfully generated to work as a matched pair in a sandwich immunoassay to detect the target protein with good sensitivity.

■ DISCUSSION

The investigation of protein structures and interactions is an essential issue for both fundamental and practical reasons. From the fundamental point of view, the development of

dx.doi.org/10.1021/cb300487u | *ACS Chem. Biol.* 2013, 8, 397−404

rational approaches to predict and design sequences with specific properties can help in understanding the physical basis of molecular recognition as well as furthering our understanding of the relationships between protein sequence, structure, dynamics, and function. From the practical point of view, designed biomolecules can have importance in the development of diverse applications, ranging from analytics and diagnostics, to drug-discovery and biotechnology. Hence, by increasing our understanding of the molecular-level origins of protein interactions, we will be able to rationally engineer novel molecules (peptides, peptidomimetics, and *de novo* designed proteins) with characteristics suitable for a particular application.

In this work, we have described a novel strategy that integrates the computational prediction of the antibody-binding regions of proteins and peptide design, with microarray technology for the development of antibody-based diagnostic tools aimed at revealing specific protein targets. The predicted epitopes were synthesized and used to induce the production of specific antibodies in rabbits. Importantly, once isolated, the elicited antibodies were shown to specifically and effectively recognize the full-length parent protein from which the epitope-mimics were derived. Competition experiments in which either of the two antibodies is coincubated with the epitope-mimics show a marked decrease or even abrogation of antibody binding, indicating a direct competition for the recognition site on the antibody. Overall, these results represent an important validation of the MLCE concept and method that we introduced in the context of epitope prediction. The computational approach could be used to screen libraries of known antigens and efficiently design a diverse yet focused, collection of epitopes. The combination of this strategy with microarrays opens up the opportunity to generate new targeted and high-throughput diagnostic platforms.

Our results indicate that the designed epitope-mimics recapitulate the physicochemical requirements that are needed to determine specific recognition of the test antigen proteins by antibodies. In this context, our designed epitope-mimics show the capability of inducing structure specific recognition of the target epitopes on the full-length proteins. One possible explanation for this result can be found in the context of the conformational selection model:[21] the epitope mimics may exist in a dynamic equilibrium between different states available on the energy landscape, whereby conformational fluctuations determine the population of structural ensembles with specific molecular recognition properties. Similar structures, sharing similar recognition profiles, can be populated by the epitope sequences in the context of the full protein. It is worth noting that antibody binding regions on antigenic proteins are most often located on surface exposed, flexible loops that can fully exploit mechanisms of conformational selection in the recognition of binding partners. In this framework, the conformational fluctuations of a specific sequence may lead to the population of the structures required to favorably engage the antigen-binding site of the antibody. This model can also provide a rationale for the observed enhanced molecular recognition of FABP by the more flexible epitope-mimic, **f-ep2-des2**: the capability of the epitope-mimic to sample a higher number of structures increases the probability to populate the conformational subset required to recognize and bind the antibody. A similar model has been proposed by Kwong and co-workers[9] who noticed that, when considering thermodynamics and antigenic recognition, the flexibility of epitopes grafted onto selected protein scaffolds correlated with immunogenicity, with more flexible epitopes generating significantly higher immune responses than their rigid homologues. From the applicative point of view, computational analysis and epitope design can have interesting applications in the larger context of molecular diagnostics, in particular, helping to expand the use of microarray technologies. Antibody-based microarrays are indeed setting new standards for molecular profiling of complex proteomes generating details of protein expression profiles in health and disease, paving the way for novel discoveries, such as improved diagnosis and prognosis of cancer patients and biomarker discovery.[22] Additional technological advances must, however, be accomplished before this tool better serves human diagnostics. One of these issues is that most antibody arrays contain a few well-defined capture agents directed at a particular class of protein markers since it is still difficult to obtain large sets of highly specific and well-validated antibody molecules. Moreover, additional antibodies may be required to allow the detection of post-translational modifications. The approach presented here represents a valuable tool to (i) exploit the analysis and prediction of the structural and recognition properties of epitopes on target protein antigens to elicit specific antibodies for improved microimmunoassays and (ii) select and focus on the sequences that are needed for the development of tailored microarrays for biomarker discovery. Biomarkers are important indicators of a certain biological state and can be used in the diagnosis of the pathogenesis of a disease, improving its treatment. So far, attempts to identify antibody biomarkers have used antigens to capture antibodies that are overproduced in disease and are present in patients' sera. In this context, combining protein microarray technology and rational design of diagnostic probes can lead to the development of microarrays displaying highly specific antigens whose properties are designed to profile patient's antibodies for diagnosing, serotyping, and monitoring infection and vaccination with impacts not only in diagnostics but also in clinical research and immunotherapy. In the larger framework of structural immunology, reliable prediction of antibody binding sites for a specific protein is an essential requirement for the selection/design of protein candidates with optimized properties in terms of sequence, structure, and presentation of the determinants for antibody-recognition. Consequently, it is important to develop methods that identify potential epitope regions of an antigen based on its structure and physicochemical properties and use this information for the development of diagnostic probes, antibodies as biomarkers or, ultimately, neutralizing antibodies able to provide protection against invading pathogens. In conclusion, it is now clear that information on the 3D structural organization and dynamics of pathogenic proteins can help define new rules for the design of effective biomarkers and immunological tools. The physicochemical structure-based approach we have presented here can be used for focusing on conformational or linear epitopes and to design immunogens able to induce specific antibodies. This detailed knowledge could in turn be used to guide the (re)design of protein antigens, to minimize their complexity, improve their physicochemical properties without impairing immune reactivity, or to broaden their efficacy. The novel integrated strategy and the results we have discussed here represent one first step toward these ambitious goals. Refinements and more tests are needed to optimize this protocol so that it can have a real impact in basic and applied molecular medicine. Given this

caveat, it is worth underlining that the aim of the study was to introduce a new integrated computational and experimental approach. Improvements from the computational side can be garnered from the dramatic improvements obtained recently in algorithms and hardware solutions,[23,24] which are starting to make it possible to obtain large scale MD-based predictions faster and on longer time scales. From the experimental point of view, advances in surface-chemistry and detection methodologies coupled to improvements in microarrays and protein production technologies have the potential to enhance the quality of the results as well as to shorten the time required to obtain them.

MATERIALS AND METHODS

Molecular Dynamics Simulations. The structures of FABP and S100B (PDB accessions are, respectively, 1G5W and 2PRU; the first structure of the NMR bundle was used in both cases) were selected as a starting point for two 30 ns all-atom MD simulations in explicit water at 300 K. Both simulations and the subsequent analysis of the trajectories were performed using the GROMACS 4.51 software package,[25] GROMOS96 force field,[26] and the SPC water model.[27]

Regarding the calculation set up used to mimic proper solution conditions: The ε-amino groups were considered protonated, while the carboxyl groups were considered to bear a negative charge. The systems were solvated in a cubic-shaped box large enough to contain 1.2 nm of solvent around the solutes. The systems were subsequently energy minimized using the steepest descent and conjugated gradients methods. The calculation of electrostatic forces utilized the PME implementation of the Ewald summation method. The constraining of all bond lengths was performed with the LINCS algorithm.[28] A dielectric permittivity, $\varepsilon = 1$, and a time step of 2 fs were used. All atoms were given an initial velocity obtained from a Maxwellian distribution at the initial temperature of 300 K. The density of the system was adjusted performing the first equilibration runs at NPT condition by weak coupling to a bath of constant pressure ($P_0 = 1$ bar; coupling time, $\tau_P = 1$ ps).[29] In both simulations, the temperature was maintained close to the intended values using Berendsen thermostat with a coupling constant of 0.2 ps.[29] The protein and the rest of the system were coupled separately to the temperature bath. The simulations were run at NPT conditions for 30 ns. The first 5 ns of each trajectory were not used in the subsequent analysis in order to minimize convergence artifacts. Equilibration of the trajectories was checked by monitoring the equilibration of quantities such as the RMSD with respect to the initial structure, internal protein energy, and fluctuations calculated on different time intervals. The detailed results of the simulations, the results of the cluster analysis, the representative structures of the main clusters, and the results of MLCE analysis are presented as Supporting Information.

Computational Epitope Prediction and Design. Epitope prediction was carried out on the representative structure of the most populated structural cluster obtained using the method developed by Daura et al.[30] The MLCE method[13] is based on the eigenvalue decomposition of the matrix of residue−residue energy couplings. For a protein of N residues, the $N \times N$ matrix (M_{ij}) of nonbonded interactions between residue pairs is built[31−36] and is then diagonalized and reexpressed in terms of eigenvalues and eigenvectors. The map of pair energy-couplings corresponding to the lowest eigenvector filtered with the contact matrix can be used to identify local couplings characterized by energetic interactions of minimal intensities.[13] Local low energy couplings identify those sites in which interaction networks are not energetically optimized. These regions may be apt to interact with binding partners or to otherwise tolerate mutations that would preserve the antigen three-dimensional structure, which is a distinctive trait of epitopes.[37] Moreover, thanks to the low intensity constraints to the rest of the protein, these substructures would be characterized by dynamic properties that allow them to visit multiple conformations, a subset of which can be recognized by the antibody to form a complex.[38−42] The lowest 15% of all contact-filtered pairs define the residue making up the predicted epitope sequences. The MLCE method is available as the free web tool BEPPE, at http://bioinf.uab.es/BEPPE.

Epitope Synthesis. The results of the structure-based prediction were used to design and synthesize a set of polypeptides corresponding to the selected epitope regions, with different functionalizations (using PEG as spacer and Keyhole Limpet Hemocyanin or Rabbit Serum Albumin (RSA) as carrier proteins)

Chemicals. Cholorotrityl chloride resin (CTC) and *N-a*-Fmoc-L-amino acids used during chain assembly were purchased from Iris Biotech GmbH (Marktredwitz, Germany). Fmoc-PEG building blocks used were 8-(9-fluorenyloxycarbonyl-amino)-3,6,dioxaoctanoic acid (Fmoc-O2Oc-OH) and 5-(9-fluorenyloxycarbonyl-amino)-3-oxapentanoic acid (Fmoc-O1Pen-OH), that were as well purchased from Iris Biotech GmbH (Marktredwitz, Germany). Sulfosuccinimidyl 4-(*N*-maleimidomethyl) cyclohexane-1-carboxylate (Sulfo-SMCC) was from Calbiochem. [2-(1*H*-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate] (HBTU) was purchased from Fluka (Buchs, Switzerland). *N*,*N*′-Dimethylformamide (DMF) and trifluoroacetic acid (TFA) were from Carlo Erba (Rodano, Italy). *N*,*N*′-Diisopropylethylamine (DIEA), dichloromethane (DCM), and all other organic reagents and solvents, unless stated otherwise, were purchased in high purity from Sigma-Aldrich (Steinheim, Germany). All solvents for solid-phase peptide synthesis (SPPS) were used without further purification. HPLC grade acetonitrile and ultrapure 18.2 Ω water (Milli-Q) were used for the preparation of all solvents for liquid chromatography.

Peptide Synthesis. All peptides were assembled manually by stepwise Fmoc-SPPS[43,44] onto a preloaded CTC resin, using HBTU/DIEA for *in situ* activation of entering amino acids. Piperidine 20% in DMF was used for Fmoc removal steps. Cysteine was added at peptides N-termini or C-termini to enable specific conjugation to carrier protein. Poorly immunogenic PEG reagents were used as solubilizing spacers to link peptide subunits. Upon completion of peptides assembling, free amino peptides were simultaneously cleaved from the resin and side chain deprotected by treatment with a mixture of 5% water, 5% phenol, 5% thioanisole, 2.5% ethanedithiol, 2.5% triisopropylsilane, and 80% trifluoroacetic acid (reagent K) for 3 h at RT. Crude peptides were precipitated and washed with cold *tert*-butylmethylether and eventually collected by centrifugation. Peptides were then dissolved in 50% aqueous ACN/0.1% TFA and purified by C_{18}-RP-HPLC. Peptide purity and identity was assessed by analytical C_{18}-RP-HPLC and following ESI-MS analysis.

Peptide Conjugation to Carrier Proteins. Peptides were specifically conjugated to freshly prepared maleimido-activated Hemocyanin from Concholepas and Rabbit Serum Albumin, using sulfosuccinimidyl 4-(*N*-maleimidomethyl) cyclohexane-1-carboxylate (sulfo-SMCC), according to the procedure described by Liu et al.[45] Polyclonal antibodies were raised in rabbits (Primm srl, Milano Italy). Antisera were immunopurified against peptide chemically linked to cyanogen bromide activated sepharose (from Sigma-Aldrich) following the instruction procedures of the manufacturer.

Antibody Production and Purification. Polyclonal antibodies were raised in rabbits by Primm srl, Milano, Italy, according to their standard immunization protocols. Antisera were immunopurified against peptides chemically linked to cyanogen bromide activated sepharose from Sigma (St. Louis, MO).

Microarray Analysis. Silicon slides (SVM, Santa Clara, CA) were coated by copoly(DMA-NAS-MAPS) as described in ref 19. Human fatty acid binding protein (FABP) was from USBiological (Swampscott, MA); S100 protein from MyBioSource (San Diego, CA); and Ovoalbumin, Ovomucoid, Rabbit, Goat, and Mouse Immunoglobulins were from Sigma (St. Louis MO). The FABP and S-100 generated antibodies were labeled with Cy3-Mono-Reactive Dye Pack from GE Healthcare (Little Chalfont, UK) according to the manufacturer's procedure. Antibody and protein microarrays were patterned on copoly(DMA-NAS-MAPS) coated silicon slides using a SciFlexArrayer S5 spotter from Scienion (Berlin, Germany) printing 1 mg/mL protein concentration in PBS. Printed slides were placed in a humid chamber and incubated at RT overnight. The slides were then blocked by 50

mM ethanolamine in TRIS/HCl, 1 M, pH 9, for 1 h, washed with water, and dried by a stream of nitrogen. In the FABP related microarray tests, arrayed slides were incubated for 2 h with FABP fluorescent antibodies 1 μg/mL in incubation buffer (TRIS/HCl, 50 mM, pH 7.6; NaCl, 150 mM; BSA, 1% w/v; Tween 20, 0.02% v/v) then rinsed by the washing buffer (Tris/HCl, 0.05 M, pH 9; NaCl, 0.25 M; Tween 20, 0.05% v/v) for 10 min under stirring and rinsed with water. For the competition experiments, FABP antibody solutions were preincubated for 60 min at 37 °C with equimolar solution of the eliciting peptides. In the S-100 immunoassays, antibody array was incubated with concentration ranging from 0 to 100 ng/mL of S-100 in incubation buffer for 2 h, then probed with fluorescently labeled detection antibody, 1 μg/mL, in incubation buffer and rinsed with washing buffer and water. Scanning for fluorescence evaluation was performed by a ProScanArray scanner from Perkin-Elmer (Boston, MA); silicon slides were analyzed using 75% and 75% of photomultiplier (PMT) gain and laser power, respectively. Fluorescence intensities of four replicated spots were averaged.

■ ASSOCIATED CONTENT

Supporting Information

Root mean square deviations of FABP and S100 simulations; MLCE predictions calculated over the representative structures of FABP and S100 after considering the clustering of the protein trajectories. This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: ++39-02-28500031. Fax: ++39-02-28901239. E-mail: g.colombo@icrm.cnr.it.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by CARIPLO "From Genome to Antigen: a Multidisciplinary Approach towards the Development of an Effective Vaccine against *Burkholderia pseudomallei,* the Etiological Agent of Melioidosis" (contract number 2009-3577). G.C. also acknowledges funding from AIRC (Associazione Italiana Ricerca sul Cancro) through the grant IG 11775 and from Fondazione Cariplo through grant 2011.1800 for the RST call "Premio fondazione cariplo per la ricerca di frontiera". G.C. thanks Asce Eastro for invaluable discussions and inspiration.

■ REFERENCES

(1) Glennie, M. J., and Johnson, P. W. (2000) Clinical trials of antibody therapy. *Immunol. Today 21*, 403−410.

(2) Morshed, S. A., Parveen, S., Leckman, J. F., Mercadante, M. T., Bittencourt Kiss, M. H., Miguel, E. C., Arman, A., Yazgan, Y., Fujii, T., Paul, S., Peterson, B. S., Zhang, H., King, R., Scahill, L., and Lombroso, P. J. (2001) Antibodies against neural, nuclear, cytoskeletal, and streptococcal epitopes in children and adults with Tourette's syndrome, Sydenham's chorea, and autoimmune disorders. *Biol. Psychiatry 50*, 566−577.

(3) Balazs, A. B., Chen, J., Hong, C. M., Rao, D. S., Yang, L., and Baltimore, D. (2012) Antibody-based protection against HIV infection by vectored immunoprophylaxis. *Nature 481*, 81−84.

(4) Mendelsohn, J., and Baselga, J. (2000) The EGF receptor family as targets for cancer therapy. *Oncogene 19*, 6550−6565.

(5) van den Elsen, J., Vandeputte-Rutten, L., Kroon, J., and Gros, P. (1999) Bactericidal antibody recognition of meningococcal PorA by induced fit: Comparison of liganded and unliganded Fab structures. *J. Biol. Chem. 274*, 1295−1501.

(6) Dormitzer, P. R., Ulmer, J. B., and Rappuoli, R. (2008) Structure-based antigen design: A strategy for next generation vaccines. *Trends Biotechnol. 26*, 659−667.

(7) Dennis, R. B. (2010) Scaffolding to build a rational vaccine design strategy. *Proc. Natl. Acad. Sci. U.S.A. 107*, 17859−17860.

(8) Montero, M., Gulzar, N., Klaric, K. A., Donald, J. E., Lepik, C., Wu, S., Tsai, S., Julien, J. P., Hessel, A. J., Wang, S., Lu, S., Burton, D. R., Pai, E. F., DeGrado, W. F., and Scott, J. K. (2012) Neutralizing epitopes in the membrane: Proximal external region of HIV-1 gp41 are influenced by the transmembrane domain and the plasma membrane. *J. Virol. 86*, 2930−2941.

(9) Ofek, G., Guenaga, F. J., Schief, W. R., Skinner, J., Baker, D., Wyatt, R., and Kwong, P. D. (2010) Elicitation of structure-specific antibodies by epitope scaffolds. *Proc. Natl. Acad. Sci. U.S.A. 107*, 17880−17887.

(10) Mihai, L. A., Correia, B. E., Ban, Y.-E. A., Carrico, C., Kalyuzhniy, O., Chen, L., Schroeter, A., Huang, P.-S., McLellan, J. S., Kwong, P. D., Baker, D., Strong, R. K., and Schief, W. R. (2011) Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science 334*, 373−376.

(11) Chiasserini, D., Parnetti, L., Andreasson, U., Zetterberg, H., Giannandrea, D., Calabresi, P., and Blennow, K. (2010) CSF levels of heart fatty acid binding protein are altered during early phases of Alzheimer's disease. *J. Alzheimer's Dis. 22*, 1281−1288.

(12) Steiner, J., Bogerts, B., Schroeter, M. L., and Bernstein, H.-G. (2011) S100B protein in neurodegenerative disorders. *Clin. Chem. Lab. Med. 49*, 409−424.

(13) Scarabelli, G., Morra, G., and Colombo, G. (2010) Predicting interaction sited from the energetics of isolated proteins: A new approach to epitope mapping. *Biophys. J. 98*, 1966−1975.

(14) Soriani, M., Petit, P., Grifantini, R., Petracca, R., Gancitano, G., Frigimelica, E., Nardelli, F., Garcia, C., Spinelli, S., Scarabelli, G., Fiorucci, S., Affentranger, R., Ferrer-Navarro, M., Zacharias, M., Colombo, G., Vuillard, L., Daura, X., and Grandi, G. (2010) Exploiting antigenic diversity for vaccine design: the chlamydia ArtJ paradigm. *J. Biol. Chem. 285*, 30126−30138.

(15) Pastori, C., Clivio, A., Diomede, L., Consonni, R., De Mori, G. M., Longhi, R., Colombo, G., and Lopalco, L. (2008) Two amino acid substitutions within the first external loop of CCR5 induce human immunodeficiency virus-blocking antibodies in mice and chickens. *J. Virol. 82*, 4125−4134.

(16) Tabatabai, L. B., and Pugh, G. W. J. (1994) Modulation of immune responses in Balb/c mice vaccinated with *Brucella abortus* Cu−Zn superoxide dismutase synthetic peptide vaccine. *Vaccine 12*, 919−924.

(17) Kwong, P. D., and Wilson, I. A. (2009) HIV-1 and influenza antibodies: Seeing antigens in new ways. *Nat. Immunol. 10*, 573−578.

(18) Kwong, P. D., Mascola, J. R., and Nabel, G. J. (2011) Rational design of vaccines to elicit broadly neutralizing antibodies to HIV-1. *Cold Spring Harb. Perspect. Med.*, a007278.

(19) Cretich, M., Di Carlo, G., Longhi, R., Gotti, C., Spinella, N., Coffa, S., Galati, C., Renna, L., and Chiari, M. (2009) High sensitivity protein assays on microarray silicon slides. *Anal. Chem. 81*, 5197−5203.

(20) Haab, B. B., Dunham, M. J., and Brown, P. O. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol. 2*, 10004.13.

(21) Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recongition. *Nat. Chem. Biol. 5*, 789−796.

(22) Borrebaek, C. A. K., and Wingren, C. (2009) Design of high-density antibody microarrays for disease proteomics: Key technological issues. *J. Proteomics 2009*, 6.

(23) Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O., and Shaw, D. E. (2009) Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol. 19*, 120−127.

(24) Shaw, D. E., Deneroff, M. M., Dror, R. O., Kuskin, J. S., Larson, R. H., Salmon, J. K., Young, C., Batson, B., Bowers, K. J., Chao, J. C.,

Eastwood, M. P., Gagliardo, J., Grossman, J. P., Ho, C. R., Ierardi, D. J., Kolossvary, I., Klepeis, J. L., Layman, T., McLeavey, C., Moraes, M. A., Mueller, R., Priest, E. C., Shan, Y. B., Spengler, J., Theobald, M., Towles, B., and Wang, S. C. (2008) Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM 51*, 91−97.
(25) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput. 4*, 435−447.
(26) van Gunsteren, W. F., Daura, X., and Mark, A. E. (1998) GROMOS force field. *Encyclopedia Comput. Chem. 2*, 1211−1216.
(27) Berendsen, H. J. C., Grigera, J. R., and Straatsma, P. R. (1987) The missing term in effective pair potentials. *J. Phys. Chem. 91*, 6269−6271.
(28) Hess, B., Bekker, H., Fraaije, J. G. E. M., and Berendsen, H. J. C. (1997) A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463−1472.
(29) Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F. V., Nola, A. D., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3690.
(30) Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W. F., and Mark, A. E. (1999) Peptide folding: When simulation meets experiment. *Angew. Chem., Int. Ed. 38*, 236−240.
(31) Colacino, S., Tiana, G., Broglia, R. A., and Colombo, G. (2006) The determinants of stability in the human prion protein: Insights into the folding and misfolding from the analysis of the change in the stabilization energy distribution in different condition. *Proteins: Struct., Funct., Bioinf. 62*, 698−707.
(32) Colacino, S., Tiana, G., and Colombo, G. (2006) Similar folds with different stabilization mechanisms: The cases of Prion and Doppel proteins. *BMC Struct. Biol. 6*, 17.
(33) Tiana, G., Simona, F., De Mori, G. M. S., Broglia, R. A., and Colombo, G. (2004) Understanding the determinants of stability and folding of small globular proteins from their energetics. *Protein Sci. 13*, 113−124.
(34) Morra, G., and Colombo, G. (2008) Relationship between energy distribution and fold stability: Insights from molecular dynamics simulations of native and mutant proteins. *Proteins: Struct., Funct., Bioinf. 72*, 660−672.
(35) Genoni, A., Morra, G., Merz, K. M. M. J., and Colombo, G. (2010) Computational study of the resistance shown by the subtype B/HIV-1 protease to currently known inhibitors. *Biochemistry 49*, 4283−4295.
(36) Torella, R., Moroni, E., Caselle, M., Morra, G., and Colombo, G. (2010) Investigating dynamic and energetic determinants of protein nucleic acid recognition: analysis of the zinc finger zif268-DNA complexes. *BMC Struct. Biol. 10*, 42.
(37) Rubinstein, N. D., Mayrose, I., Halperin, D., Yekutieli, D., Gershoni, J. M., and Pupko, T. (2008) Computational characterization of B-cell epitopes. *Mol. Immunol. 45*, 3477−3489.
(38) Ma, B. Y., Wolfson, H. J., and Nussinov, R. (2001) Protein functional epitopes: Hot spots, dynamics and combinatorial libraries. *Curr. Opin. Struct. Biol. 11*, 364−369.
(39) Ponomarenko, J. V., and Bourne, P. E. (2007) Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *BMC Struct. Biol. 7*, 64.
(40) Zhang, Q., Wang, P., Kim, Y., Haste-Andersen, P., Beaver, J., Bourne, P. E., Bui, H. H., Buus, S., Frankild, S., Greenbaum, J., Lund, O., Lundegaard, C., Nielsen, M., Ponomarenko, J., Sette, A., Zhu, Z., and Peters, B. (2008) Immune epitope database analysis resource (IEDB-AR). *Nucleic Acids Res. 36*, W513−W518.
(41) Lichtarge, O., Bourne, H. R., and Cohen, F. E. (1996) An evolutionary trace method defines binding surfaces common to protein families. *J. Mol. Biol. 257*, 342−358.
(42) Westhof, E., Altschuh, D., Moras, D., Bloomer, A. C., Mondragon, A., Klug, A., and Van Regenmortel, M. H. (1984) Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature 311*, 123−126.
(43) Atherton, E., and Sheppard, R. C. (1987) in *The Peptides* (Udenfriend, S., and Meienhofer, J., Eds.) Vol. *9*, pp 1−39, Academic Press, San Diego, CA.
(44) Fields, G. B., and Noble, R. L. (1990) Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res 35*, 161−214.
(45) Liu, Z., Gurlo, T., and von Grafenstein, H. (2000) Cell-ELISA using β-galactosidase conjugated antibodies. *J. Immunol. Methods 234*, 153−167.