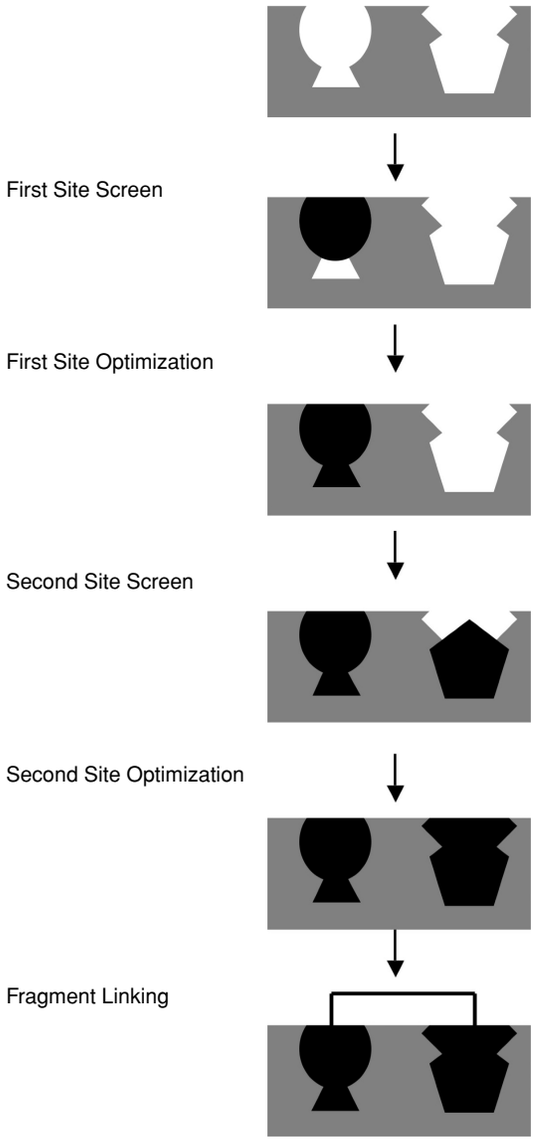

Figure 4.3. Outline of the SAR by NMR method. The SAR by NMR method consists of several steps. First, a molecule is identified that binds to one subsite. This scaffold is then optimized for affinity to the first subsite. A second compound is then identified that binds to an adjacent subsite. The second ligand is then optimized for maximum affinity. Finally, based on structural input, a linker is designed to optimally connect the two scaffolds (adapted from Shuker et al.[1]).

Direct detection of ligand binding can be a powerful technique in providing information for the drug discovery process. Because the sequence-specific assignments of the spectrum are usually determined before screening takes place, localization of ligand binding is revealed by the observed chemical shift perturbations. In addition, this information allows the discrimination of nonspecific and nonrelevant ligand binding from binding that affects the activity of the target protein and can be used as input to processes to develop more potent binding ligands, such as SAR by NMR.[1] This approach (illustrated in Figure 4.3) relies on the identification of two distinct fragment binding sites on the target protein that are close in proximity. Optimal linking of the two fragments may result in a high-affinity ligand. SAR by NMR has become the principal motivation for target-directed detection of fragment binding using uniformly labeled proteins. Numerous applications of SAR by NMR and similar target-directed strategies have been reviewed in the literature.[11,49–56]

The use of chemical shift perturbation in uniformly labeled proteins to detect binding is limited by several factors. Because many ^{15}N-^{1}H correlation spectra are required to both screen the library and deconvolute the binding mixtures, a relatively large quantity of ^{15}N labeled protein must be available before studies can be undertaken. Furthermore, the target of interest must be sufficiently soluble and present in high concentrations so that data can be acquired without the need for excessive signal averaging. The actual quantity of protein required will depend on the library size and composition as well as the sensitivity of the available NMR instrumentation. For example, if reasonable data can be acquired at a target protein concentration of 100 μM on a 25-kDa protein, then screening 200 NMR samples of a 500-μl volume would require 250 mg of labeled protein. This would be sufficient to screen a library of 1,000–2,000 compounds (neglecting deconvolution) if each mixture contained five to ten fragments. Additional labeled protein would be required to obtain spectra of individual compounds from the "hit" mixtures to identify the binding ligand. For these reasons, some groups have developed strategies to minimize protein consumption and labeling costs, such as using NMR flow systems.[57]

To take full advantage of the information provided by target-directed screening methods, it is advantageous to have the sequence-specific resonance assignments of the target protein available. This requirement imposes further limitations on the method. For well-behaved, highly soluble proteins with molecular masses below approximately 25 kDa, three-dimensional triple resonance spectra of doubly labeled proteins can often be used efficiently to provide these assignments. At higher molecular masses, the requirement to perdeuterate the target protein to acquire triple resonance spectra[58] creates additional difficulties in protein expression. The lengthy data acquisition process, combined with complexity of obtaining resonance assignments of these large proteins, may make the process so time-consuming as to reduce the impact of information generated by NMR detected fragment screening approaches.

One technique to circumvent many of the limitations associated with uniformly labeled target protein has been described previously.[59] Provided that the binding site of interest is known independently (for example, by mutagenesis or structural studies) and it contains a sequentially unique pair of amino acids, an amino acid–specific labeling strategy may be employed to detect relevant binding without the need for determining complete sequence-specific resonance assignments. In this method, the *i*th