J Comput Aided Mol Des (2009) 23:475–489
DOI 10.1007/s10822-009-9287-y

The multi-copy simultaneous search methodology: a fundamental tool for structure-based drug design

Christian R. Schubert · Collin M. Stultz

Received: 1 January 2009 / Accepted: 20 May 2009 / Published online: 9 June 2009

Abstract Fragment-based ligand design approaches, such as the multi-copy simultaneous search (MCSS) methodology, have proven to be useful tools in the search for novel therapeutic compounds that bind pre-specified targets of known structure. MCSS offers a variety of advantages over more traditional high-throughput screening methods, and has been applied successfully to challenging targets. The methodology is quite general and can be used to construct functionality maps for proteins, DNA, and RNA. In this review, we describe the main aspects of the MCSS method and outline the general use of the methodology as a fundamental tool to guide the design of de novo lead compounds. We focus our discussion on the evaluation of MCSS results and the incorporation of protein flexibility into the methodology. In addition, we demonstrate on several specific examples how the information arising from the MCSS functionality maps has been successfully used to predict ligand binding to protein targets and RNA.

C. R. Schubert
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

C. R. Schubert · C. M. Stultz (✉)
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
e-mail: cmstultz@mit.edu

C. M. Stultz
Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

C. M. Stultz
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Keywords Structure-based ligand design · Multi-copy simultaneous search (MCSS) · Ligand binding · Molecular drug design

Introduction

After nearly three decades since its theoretical inception [1], rational drug design remains a challenge. Typically, the design of compounds that bind pre-specified targets of known structure is one of the first steps in the construction of novel therapeutic compounds [1]. Methods that simplify the discovery of such *lead compounds* can accelerate the rate at which new drugs are discovered. The discovery of ligands that bind targets with known structure can be approached using a series of steps that include: (1) determining optimal positions of small chemical fragments in the binding site; (2) linking the fragments together to form molecules that are complementary to the target; (3) estimating the binding affinity of the resulting molecules; and (4) synthesis and experimental evaluation of the computationally designed ligands (Fig. 1). The multi-copy simultaneous search (MCSS) methodology [2] addresses the first step in the above-mentioned approach: the MCSS method finds energetically favorable positions of different functional groups in a pre-specified binding site of interest, and yields so-called functionality maps of the binding site as output. Functional groups positioned in these energetically favorable positions can then be linked together to construct ligands de novo that are complementary to the binding site of the target. Alternatively, these positions can be used to modify known ligands to improve their binding affinity. The method is quite general and can be used to construct functionality maps for proteins, DNA, and RNA. Moreover, protein flexibility can be incorporated in the process in a straightforward manner.

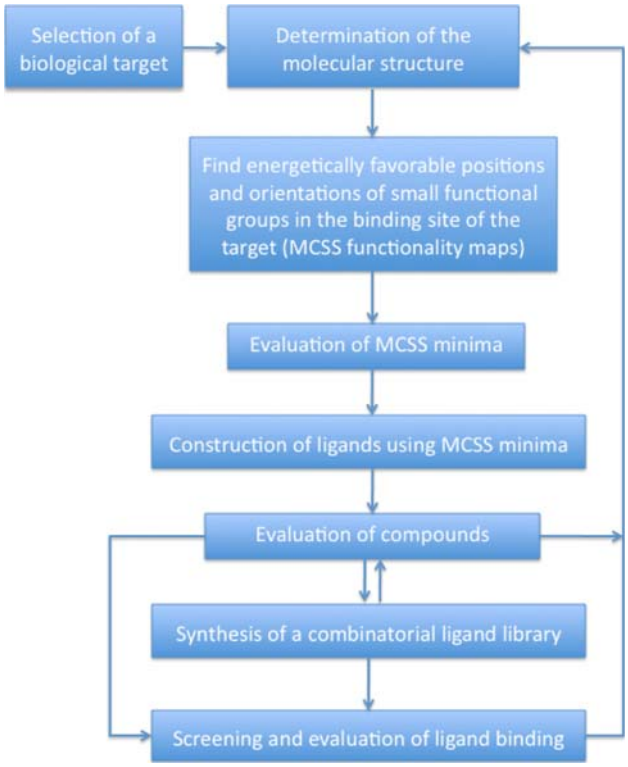

Fig. 1 An overview of a general ligand discovery strategy using MCSS. If the structure of the biological target molecule is known, the multiple copy simultaneous search methodology (MCSS) can be used to identify structural fragments that bind preferentially to the target molecule. The energetic analysis and evaluation of MCSS results can subsequently be used to propose a set of ligands that can be synthesized and screened in biological assays

In this review, we build upon our previous work that outlines the MCSS method [3]. We describe the main aspects of the methodology with a focus on the incorporation of protein flexibility and MCSS minima evaluation, and demonstrate on several specific examples how the information arising from the functionality maps has been successfully used to predict ligand binding to proteins and RNA. We outline how the methodology can be used as a fundamental tool to design novel therapeutic compounds.

The MCSS methodology

First, we review the major aspects of the general MCSS methodology and the choice of functional groups used for the calculation of functionality maps, and lastly illustrate the method on an example protein fixed as a rigid target. This section is later followed by a more focused and in depth discussion on the problem of careful evaluation of functional group maps and incorporation of protein flexibility into the MCSS methodology.

MCSS minimizations

The MCSS methodology begins by randomly distributing many replicas, usually between 1,000 and 10,000 copies, of a specific small functional group in the binding site of a target molecule. The positions of the different functional group copies are then simultaneously minimized in the binding site using cycles of conjugant gradient minimization. To save time, functional group replicas that are converging to a common minimum energy position are removed during the minimization process. In most applications, replicas are said to be converging to the same minimum energy position if their root mean square deviation (rmsd) is less than 0.2 Å. Use of such a small cutoff ensures that a variety of different functional group positions and orientations will be retained. The general methodology is outlined in Fig. 2.

A central component of the method is the time-dependent Hartree (TDH) approximation, which enables a number of molecular trajectories to be simultaneously determined in the force field of a target molecule [4, 5]. Under these conditions, different functional group replicas do not see each other, but do see the full force field of the target molecule. In a number of applications, the functional group replicas are fully flexible and therefore sample different conformations during the minimization procedure, while the target molecule is fixed. When the target is rigid, the trajectories generated with the TDH approximation are exact. In other words, the MCSS minimum energy positions are identical to minima that would have been obtained if each randomly placed replica were individually minimized in the force field of the fixed protein.

By contrast, if the protein is flexible, the resulting MCSS minima may differ from positions obtained from individual minimizations [6, 7]. Many of the initial applications of MCSS focused on obtaining functionality maps to rigid targets. However, protein flexibility can be incorporated into MCSS minimizations in a straightforward manner [8], and the resulting data can be post-processed to determine the minima that are most promising. As mentioned earlier, the issue of inclusion of protein flexibility into MCSS calculations is discussed in depth in a later section of the text.

Choice of functional groups

The chemical fragments used during MCSS calculations are typically small functionalities spanning the range of different chemical moieties that could be used to build larger, more realistic molecules. When MCSS was first introduced, calculations were performed with functional groups that typically contained less than 8 atoms [2], but essentially complete ligand candidates can be used today.

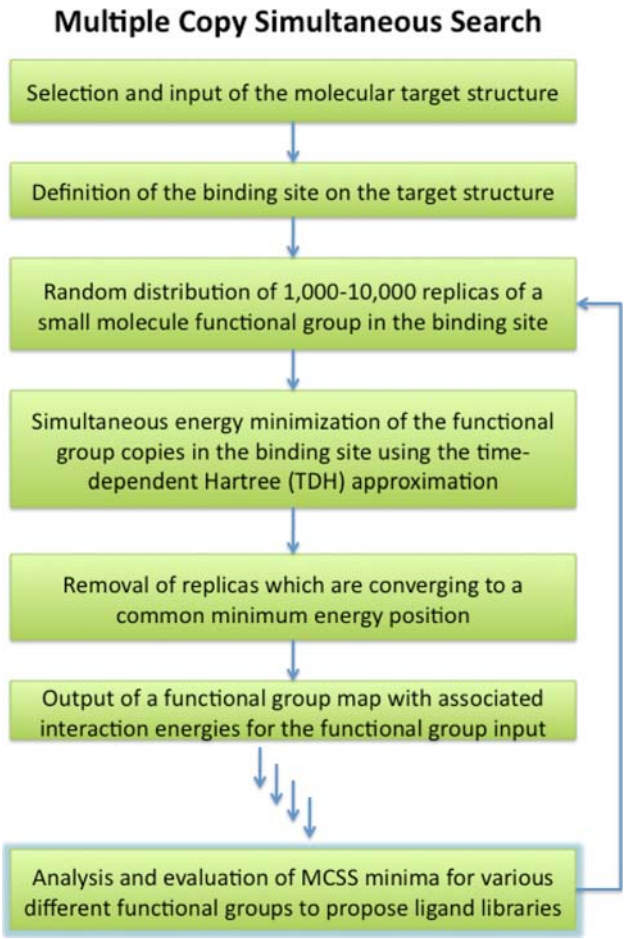

Fig. 2 A flowchart of the standard MCSS methodology. After input of the target structure and addition of hydrogen atoms a binding site in form of a sphere or a box is defined, and usually between 1,000 and 10,000 copies of a specified functional group are randomly distributed within it. The positions of the functional groups are simultaneously minimized using the time-dependent Hartree (TDH) approximation, so that different functional group replicas do not see one another, but see the full force field of the target molecule. Functional group replicas converging to the same minimum energy position, are removed during the minimization process to save time. Once the functional group output map is generated, the MCSS minima can be analyzed and evaluated, and if applicable, the process can be repeated. Ultimately, the MCSS fragment maps for various different functional groups can be combined to propose small ligand libraries

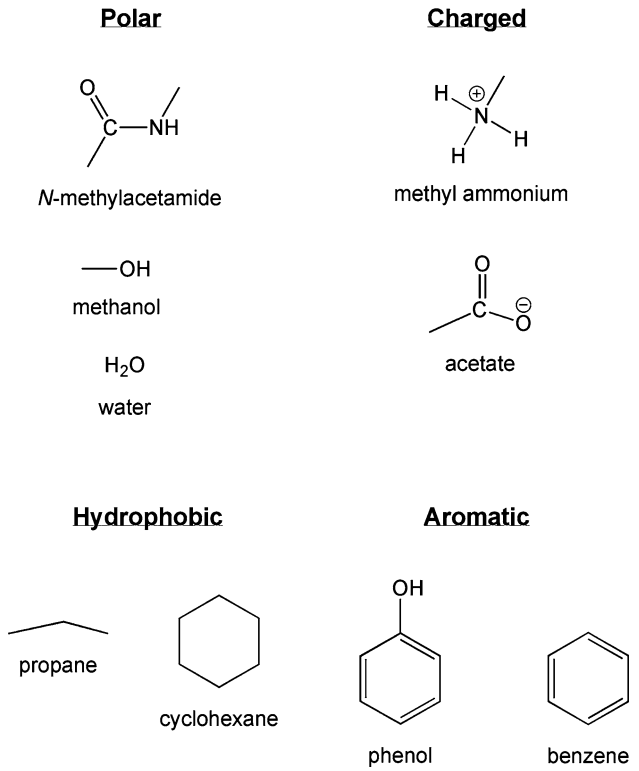

Fig. 3 Popular functional groups used in MCSS minimizations. These groups collectively model polar, charged, hydrophobic, and aromatic moieties found in many biological ligands. For a more extensive review of available chemical moieties that have been incorporated into the standard MCSS methodology, see Stultz and Karplus (2006) [3]

Popular functional groups used in prior applications include *N*-methylacetamide, methanol, methyl ammonium, acetate, propane, benzene, phenol, cyclohexane, and water [2, 9–11]. Collectively, these groups model polar, charged, and hydrophobic moieties of different sizes (Fig. 3). *N*-methylacetamide (NMA) is chosen to model the peptide backbone, and is of particular interest if the goal is to design peptide-based ligands that are complementary to a given target protein.

As computational performance and speed have improved dramatically over the last two decades, the size of functional groups used in routine MCSS calculations has contemporaneously increased significantly: MCSS calculations can now be performed on rather complex chemical fragments that contain 20–30 atoms [12]. Functional groups currently available with the standard version of MCSS span a range of possible chemical fragments and properties [3], including fragments that were obtained from an analysis of the structures found in a database of known drug molecules [13, 14].

The addition of novel or more complex functional groups to the MCSS library is relatively straightforward: it involves the creation of the necessary topology, parameter, and symmetry files. Just as in the case of the standard CHARMM topology and parameter files [15], the MCSS topology file lists the partial charges, atoms types, and connectivity of the atoms. The group parameter file contains the additional parameters that are needed for the energy calculations. MCSS runs can be performed with both a polar hydrogen parameter set [16] and an all-atom parameter set [17], or using a hybrid approach where the protein is represented with one particular parameter set, such as all-atom, and the MCSS functionalities are represented with another parameter set, such as polar.

If the chosen functional group has several rotatable bonds, it must be ensured that enough randomly placed functional group replicas are included in the initial stages of the simulation to sample the different conformations of the chemical fragment, since undersampling of functional group conformations will complicate the finding of functional group minima with a variety of different conformations.

Practical example: MCSS functionality maps of endothiapepsin

To illustrate how the aforementioned concepts of the MCSS methodology can be implemented in practice, we briefly discuss an early application of the method where functionality maps were created for the fungal protease *endothiapepsin* [18]. The structure of endothiapepsin bound to the peptide-based inhibitor H-261 is shown in Fig. 4a. The binding site of the enzyme is relatively open and contains several hydrophobic pockets that form binding sites for hydrophobic moieties within the inhibitor [19]. MCSS minimizations were performed on the binding site of endothiapepsin with the functional group probes *N*-methylacetamide (NMA), methyl ammonium, methanol and propane. Polar minima for NMA and methanol were found throughout the binding site forming hydrogen bonds to polar moieties within the protein (Fig. 4b, c). As expected, apolar functionalities are found throughout the binding site within apolar pockets (Fig. 4d), and charged groups such as methylammonium form salt bridges with complementary charged side chains in the binding site and therefore occur in smaller numbers (Fig. 4e). In addition, charged minima are located on the surface of the protein in regions relatively distant from the binding site. This comes at no surprise, since charged groups tend to be solvent-exposed at the surface of a protein in general.

Comparison of the MCSS functionality maps with the positions of corresponding functionalities in the inhibitor demonstrates that the MCSS methodology can reproduce the positions and orientations of chemical fragments found in the inhibitor. Since inhibitor H-261 is peptide-based, minima of the hydrophobic propane functional group are found in hydrophobic pockets occupied by phenylalanine, leucine and valine side chains. Similarly, minima for the hydrophilic functional group methanol are found near the lone hydroxyl moiety in H-261 that forms hydrogen bonds with two catalytic aspartic acid residues in the binding site. Lastly, most peptide bonds within H-261 are located near an NMA minimum modeling the characteristics of the peptide backbone. However, some peptide bonds within the inhibitor are not located near an MCSS minimum, and the nearest NMA minima show large root-mean-square deviations (rmsd) from the peptide bonds of the inhibitor [20]. This suggests that some peptide bonds in H-261 in the bound crystal form may not be in an energy minimum orientation, and it exemplifies how MCSS-based functional group minima could be used to improve the binding of inhibitor H-261 to endothiapepsin.

Limitations of the MCSS methodology

The MCSS methodology determines positions for small fragments and functionalities and not complete inhibitors. As mentioned before, undersampling can be a problem for large functional group fragments that include multiple rotatable bonds, as it will complicate the finding of functional group minima for a variety of different conformations. Similarly, discrepancies may be observed between the functional group positions in co-crystal structures of bound inhibitors and the positions of MCSS minima, as described above for the case study of endothiapepsin[8]. However, such observations do not come as a surprise. On the contrary, the real value in the MCSS methodology lies in the fact that MCSS typically finds a number of new minima that were not observed in known crystallographic structures, suggesting that the affinity of these inhibitors may be improved if low energy functionalities are incorporated into the inhibitor structures.

Evaluation of MCSS minima

Despite the fact that MCSS minimizations begin with several thousand randomly placed replicas within the binding site of the target molecule, typically only a few hundred replicas remain after repeated cycles of energy minimization. Each minima is associated with two specific parameters: (a) a precise geometrical position within the binding site, and (b) an associated interaction energy, representing a sum of the intragroup energy and the group-protein interaction energy [2]. Adequate post-processing techniques to evaluate MCSS minima can greatly facilitate the finding of new ligands.

Several approaches to evaluate MCSS minima have been reported, principally spanning the range from simple in vacuo calculations to remove minima with relatively unfavorable interaction energies [2, 20–22], to more elaborate methods of estimating the binding free energy of each minimum and incorporation of solvation by using implicit solvent models [23, 24].

The most extensively used method for post-processing MCSS minima is to remove minima with interaction energies above some cutoff value that is based on the solvation enthalpy of the functional group in question [2, 20–22]. This provides an approximate way to account for the fact that functional groups within the binding site are at least partially desolvated, and typically, only

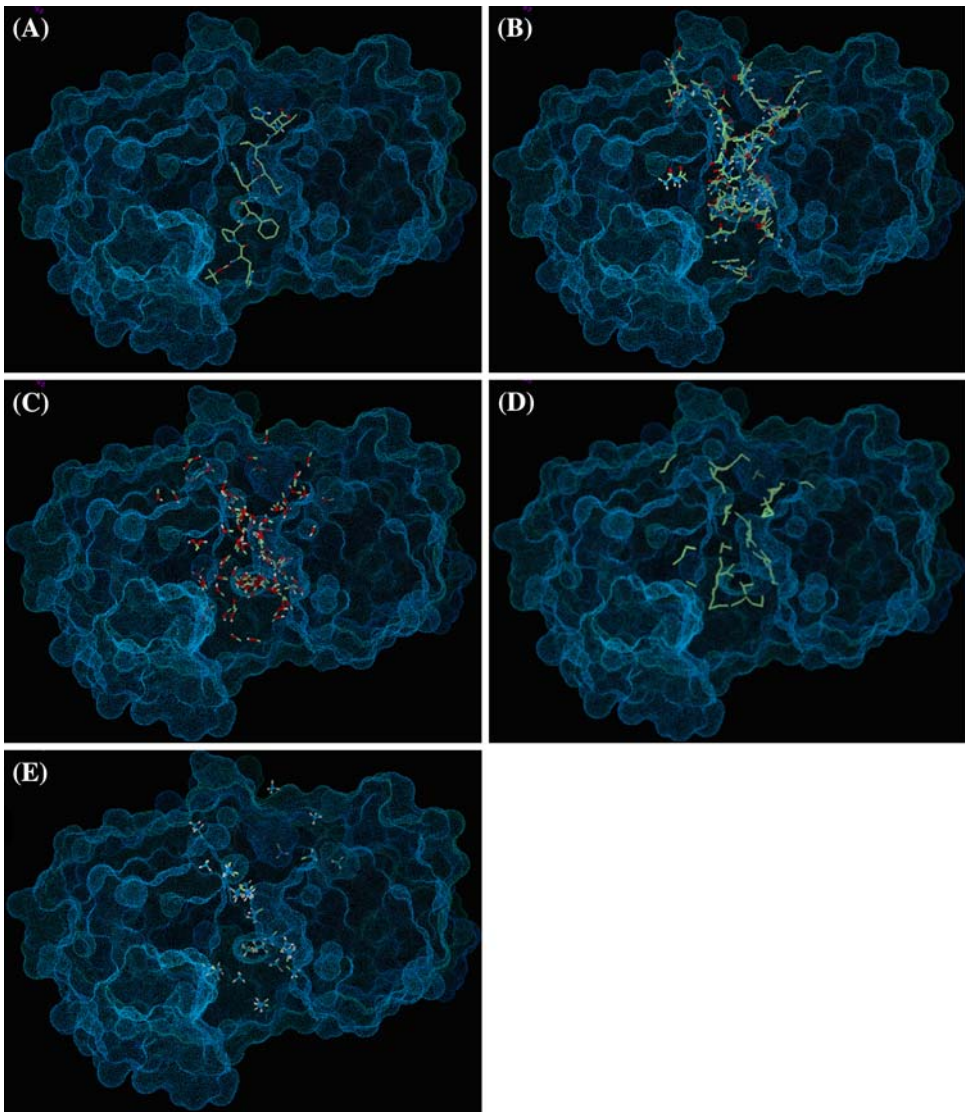

Fig. 4 MCSS functional group minima in the binding site of endothiapepsin. **a** Structure of H-261 in the endothiapepsin binding site (for comparison). **b** NMA **c** Methanol **d** Propane **e** Methyl ammonium. Reprinted with permission from Stultz and Karplus [3]

functional groups with interaction energies below the solvation energy are kept for further evaluation [3]. It is possible to conduct the minimizations with a potential energy function that incorporates an implicit solvent for model. However, as has been noted, many implicit solvent models overestimate the strength of electrostatic interactions between charged groups and hence considerable error can arise when such models are used [25, 26].

In one study, Caflisch [23] employed an alternate approach that entailed ranking minima arising from a MCSS calculation on α-thrombin by estimating the binding free energy of each minimum with an approximate free energy formula given by:

$$\Delta G_{\text{binding}} \equiv \Delta E_{\text{intra}} + \Delta E_{\text{inter,vdw}} + \Delta G^{\text{inter}}_{\text{elec}} + \Delta G^{p}_{\text{desolv}} + k\Delta G^{m}_{\text{desolv}} + \Delta G_{\text{np}} \quad (1)$$

where ΔE_{intra} denotes the intramolecular energy of the MCSS chemical fragment and is the sum of bonded, vdw, and electrostatic contributions. The intermolecular vdw interaction energy between the protein and the functional group is included in an additional term denoted by $\Delta E_{\text{inter,vdw}}$. The electrostatic interaction energy between the protein and functional group is denoted by $\Delta G^{\text{inter}}_{\text{elec}}$. The energy required to desolvate the binding site of the protein is given by $\Delta G^{p}_{\text{desolv}}$ and the energy required to desolvated the surface of the functional group that interacts with the protein is denoted by $\Delta G^{m}_{\text{desolv}}$. In this implementation, the desolvation energy of the functional group was multiplied by an additional factor, $k < 1$, which serves as an approximate way to account for the fact the desolvation contribution for a functional group which is part of a larger ligand may be smaller. This is likely to be true for the majority of small functional groups that are considered by MCSS because in the unbound state small functional groups are surrounded by other atoms in the ligand which prevent the group from being completely solvated. In the

work on α-thrombin, a value of $k = 0.4$ was used. This value was derived from a comparison of the electrostatic desolvation energy upon binding of NMA and a glycine containing dipeptide to a macromolecular target [23].

The electrostatic free energy terms were computed with a continuum dielectric model based on the linearized Poisson Boltzmann Equation and numerical solutions to the linearized Poisson–Boltzmann equation were obtained with the program UHBD [27]. The nonpolar contribution to the binding free energy, ΔG_{np}, is assumed to be linearly related to the loss in solvent-accessible surface area—a common approach for estimating the hydrophobic contribution to binding [28].

The approximate binding free energy formula ignores some contributions which are, in general, important for ligand binding. Most notably, the loss of configurational entropy upon binding is not included, even in an approximate way. This contribution is likely to be negligible for the small, relatively rigid, functional groups considered in the initial study. However, if more flexible functional groups are used, then additional terms will be needed to approximate the change in configurational entropy upon binding.

For α-thrombin, application of Eq. 1 to the MCSS minima leads to considerable reordering of the minima relative to the ordering obtained with a vacuum potential [23]. Differences between the vacuum energy rankings and the rankings obtained with Eq. 1 are primarily due to the desolvation terms. In many cases, apolar functional groups with favorable vacuum interaction energies have unfavorable binding free energies because they partially desolvate charged groups on the protein. Similarly, a number of charged groups with favorable vacuum energies have unfavorable binding free energies, primarily because of the large desolvation penalty associated with the binding of charged moieties. The inclusion of desolvation terms has the greatest affect in determining the relative affinities of minima for different functional groups. However, any potential benefit from such calculations must be weighed against the increase in computational time that is incurred with such an analysis. We note that an alternate approach would be to use the MCSS minima, and their vacuum interaction energies, to create potential ligands and then evaluate the binding energy of the resulting ligands with more rigorous approaches.

MCSS and experiment

Reproduction of functionality maps of known inhibitors

The MCSS methodology was initially tested on a small subset of functional groups that were used to map the binding site of the influenza coat protein, hemagglutinin [2]. It was demonstrated that MCSS could reproduce the positions of the corresponding functionalities of a sialic acid molecule, the natural substrate, bound to hemagglutinin. Accuracy of the method was further demonstrated by the small functional group mapping of the binding site of the HIV-1 protease and comparison to the corresponding functionalities in the known HIV-1 inhibitor MVT-101 [9]. Subsequent comparisons of MCSS minima to the structures of known inhibitors have been used to validate the MCSS approach in a number of applications, e.g. [9, 29, 30].

Reproduction of experimentally determined functionality maps of small organic molecules

In the mid 1990s, an experimental approach was developed for the finding of binding sites of small organic molecules within a target molecule of known three-dimensional structure [31]: the multiple solvent crystal structure (MSCS) method. In this approach, the crystal structure of a given protein is determined in the presence of different organic solvents, so that the resulting three-dimensional structures reveal positions of organic solvent molecules on the surface of the protein. The positions arising from the method can be directly compared to the MCSS minimum energy positions obtained when the same organic solvent is used as a functional group probe.

In the initial publication, the results of the MSCS method applied to porcine pancreatic elastase in acetonitrile were compared to MCSS minima obtained by minimizing randomly placed chemical fragments using a vacuum potential and post-processing the data to account for the effects of aqueous solvent. There was qualitative agreement between the MSCS solvent positions and the minima identified by MCSS. In another application, MSCS was applied to thermolysin in the presence of the organic solvents acetone, acetonitrile, and phenol [32]. The results were compared to MCSS functional group maps, and showed that MCSS minimum energy positions in most cases correlated well with MSCS positions of the solvent molecules. These data again suggest that MCSS can be used to effectively map binding positions of small functional groups.

However, it should be noted that experiments with organic solvent introduce significant alterations in the structure of a protein; in fact, in the case of porcine pancreatic elastase, the protein had to be cross-linked prior to the structure acquisition in organic solvent. Further, it is likely that the observed binding sites of organic molecules in the presence of non-aqueous solvent will differ significantly due to the difference in dielectric constants observed for neat organic solvents and water. Hence, comparisons between the results of MSCS experiments and MCSS calculations should be interpreted cautiously.

A complementary computational method to both MSCS and MCSS is computational solvent mapping, or CS-mapping [33]. This method determines ligand binding sites computationally by mapping the binding of small organic molecules as probes on the surface of a protein, resulting in consensus sites that bind a number of different probes with the lowest average free energy [34, 35]. First applied to hen egg-white lysozyme [36] and thermolysin [32, 37], CS-mapping yielded results in good agreement with the results of experimental mapping by MSCS. CS-mapping provides a potentially useful strategy to complement MCSS for determining ligand binding sites without the experimental pitfalls of MSCS, especially when MCSS and MSCS yield diverging results.

Employing similar tactics as MSCS, but utilizing a different experimental approach, NMR has been used to experimentally determine binding sites for small organic molecules within the binding site of a given target, and the results have been compared to functionality maps obtained from MCSS calculations [38, 39]. The underlying principle of the methodology is that variations in chemical shifts in pre-specified regions of a ^{15}N-labeled target protein are observed upon ligand binding to the region of interest, and as such, these shifts can be used to identify potential ligands from a database of small organic compounds. For example, binding sites on FKBP12 have been identified and compared to the corresponding MCSS functionality maps [39]. The MCSS functionality maps contained at least one minimum energy position that satisfied the NOE restraints arising from NMR data. Moreover, each minimum that satisfies the NOE restraints was found to be among the top ten minima in the corresponding MCSS functionality map, as estimated from an approximate binding free energy function [23, 38]. Taken together, these data strongly support the use of the MCSS methodology for the discovery of binding sites of small molecules as well as for fragments of larger compounds.

Ligand design with MCSS

In this section, we provide a brief overview of how the MCSS methodology has been used to calculate functionality maps for specific macromolecules, and to illustrate how this knowledge has facilitated the design of ligands of various types that bind with high affinity to its specific targets. A selected list of the published literature is presented in Table 1.

While the methodology in principle can be applied in a straightforward manner to RNA and DNA, the focus has remained on the design of ligands to proteins (Table 1). However, two examples can be found in the literature describing the application of the MCSS methodology to two specific types of RNA [40]. In the remainder of this section, we introduce several selected examples to illustrate how MCSS has been used successfully to design ligands to a variety of targets based on functional group maps.

Peptide-based protein ligands

In an interesting study, the MCSS methodology was used to design peptides that inhibit the association of the Ras-Raf complex [11, 41] which is assumed to play an important role in promoting general cell growth and oncogenesis [42]. The idea underlying the design strategy was to obtain a peptide that would bind at the Ras-Raf interface and prevent formation of the complex [11]. While the MCSS calculations identified three potential binding sites mimicking regions observed in crystal structures of a homologous protein bound to Raf [41], more importantly, novel interactions were identified that could be utilized in the design of a peptide based inhibitor. The approach taken is very similar to the one described earlier for the construction of peptide based ligands to the HIV-1 protease [4]: (1) energetically favorable positions for NMA, benzene, phenol, methanol, acetate, methylammonium, methylguanidium, propane were obtained using MCSS; (2) a polyglycine peptide backbone was constructed from the NMA minimum energy positions; (3) the functional group fragments mimicking specific amino acids were connected to the constructed backbone; (4) the resulting complete peptides were aligned to determine the probabilities of different side chains occupying each C_α position.

This approach yielded a total of 104 peptides, which were aligned to construct a consensus sequence. Interestingly, the consensus sequence was similar to known effector sequences that modulate the binding of Ras to Raf. These lead sequences were further refined taking advantage of sequence motifs that are known to bind Ras, and optimized for solubility. The resulting peptides were synthesized, and evaluated experimentally: they were able to inhibit the Ras–Raf interaction in vitro by 23–39% using an ELISA-based assay [11].

While the details of the described approach are in principle similar to previously described MCSS methods, there are some important differences (i.e., with regard to minima evaluation, treatment of solvation, initial undersampling). A detailed discussion would be beyond the scope of this review, however, and can be found elsewhere [3, 11, 41]. Overall, this method presents an interesting example of how information from MCSS and experimental methods can be combined to obtain peptides that inhibit the association of two proteins.

Table 1 Multiple copy simultaneous search (MCSS) in practice

Protein	Ligand	Reference
α1d Adrenergic receptor	Small molecule	*J. Med. Chem.* (2004) 47:1900
Acetylcholinesterase	Small molecule inhibitor	*Med. Chem. Res.* (1999) 9:98
Acyl carrier protein Synthase	Coenzyme A	*Proteins* (2003) 51:189
AmpC β-lactamase	Boronic acids	*J. Med. Chem.* (2002) 45:3222
Bcl-2	Small molecule inhibitor	*Bioorg. Med. Chem.* (2007) 15:6407
C. albicans N-myristoyltransferase	Small molecule inhibitor	*Huaxue Xuebao* (2008) 29:117
CYP51	Small molecule inhibitor	*J. Med. Chem.* (2003) 46:474
DHFR	NADPH/Methotrexate	*Proteins* (2003) 51:189
DNA gyrase B	DNA	*J. Med. Chem.* (2004) 47:4374
Endothiapepsin	Small functional groups	*Proteins* (2000) 40:258
Factor Xa	Small molecule inhibitor	*J. Mol. Graph. Model.* (2003) 22:105
Glycogen phosphorylase	Small molecule inhibitor	*J. Comput. Aided Mol. Des.* (2001) 15:613
Hemagglutinin	Small functional groups	*Proteins* (1991) 11:29
		Proteins (1994) 19:199
Hen egg-white lysozyme	Organic solvents	*222nd ACS Natl. Meet.* (2001) COMP-031
Hepatitis delta antigen	D-peptide inhibitors	*J. Comput. Aided Mol. Des.* (2000) 14:705
Herpes simplex polymerase	Small molecule inhibitor	*224th ACS Natl. Meet.* (2002) MEDI-021
HIV-1 aspartic proteinase	Small molecule inhibitor, small functional groups	*Proteins* (2003) 51:189
		J. Comput. Aided Mol. Des. (2000) 14:161
		Proteins (1999) 37:512
HLA-DR4	Small molecule inhibitor	*J. Comput. Aided Mol. Des.* (2007) 21:395
Human α-thrombin	Small molecule inhibitor	*J. Comput. Aided Mol. Des.* (1996) 10:372
		J. Comput. Aided Mol. Des. (1996) 10:1
Insulin	Glucose	*Proteins* (2004) 55:568
MIF	Hydroxyphenylpyruvate	*Proteins* (2003) 51:189
Peptide deformylase	Natural substrate	*217th ACS Natl. Meet.* (1999) CINF-046
Picornavirus capsid	Small molecule inhibitor	*J. Am. Chem. Soc.* (2001) 123:12758
Poliovirus capsid	Small molecule inhibitor	*J. Comput. Aided Mol. Des.* (2001) 15:935
Prostate specific antigen	Small molecule inhibitor	*J. Med. Chem.* (2001) 44:1491
		Bioorg. Med. Chem. Lett. (1997) 7:1689
Ras	Peptide	*Protein Eng.* (1999) 12:457
Ras-Raf complex	Peptide	*Protein Eng.* (2001) 14:39
Rhinovirus	Small molecule inhibitor	*214th ACS Natl. Meet.* (1997) COMP-229
RNAse A	Formate, water	*Protein Eng.* (1996) 9:773
Src SH3	Small functional groups	*J. Comput. Aided Mol. Des.* (2001) 15:935
Thermolysine	Organic solvents	*Protein Eng.* (2001) 14:47
Tyrosine kinase 2	Small molecule inhibitor	*J. Comput. Aided Mol. Des.* (2002) 16:831
VEGFR2	Small molecule inhibitor	*Huaxue Xuebao* (2007) 65:547
RNA	**Ligand**	**Reference**
16S rRNA	Paromomycin	*Theor. Chem. Acc.* (1999) 101:131
HIV-1 TAR	Arginine	*Theor. Chem. Acc.* (1999) 101:131

The references were obtained through a SciFinder Scholar search (https://scifinder.cas.org/) last accessed on November 24, 2008. The literature was searched for the keywords "multiple copy simultaneous search" and "MCSS", with further refinement for the keywords "ligand" and "ligand binding". Selected references were compared and combined in Table 1

Non-peptide based protein ligands

In another study, the MCSS methodology was used to design non-peptide based inhibitors to a member of the cytochrome P450 family, lanosterol 14α-demethylase (CYP51) [10], a proven antifungal agent. The calculations began with structural models of CYP51 from *Mycobacterium* and *Candida albicans*, and were conducted on the substrate-binding pocket using the functional groups benzene, propane, butane, cyclohexane, phenol, methanol, ethanol, ether and water. The minimizations yielded a large number of functional group minima for both protein models. For each functional group considered, the lowest energy minimum in its site was chosen for further analysis, which retained a total of only 13 minima. These were connected manually to obtain putative lead compounds, paying attention not to introduce internal strain into the structure, and assuring that the resulting molecules remained accessible by chemical synthesis.

In this study, four promising compounds were synthesized and evaluated experimentally. These compounds showed IC_{50}s in the μM range against reconstituted CYP51 from *Candida albicans*, and moreover, calculated CHARMM interaction energies that correlated quite well with the measured IC_{50}s. Modeling of the potential leads into the CYP51 binding site suggested some additional modifications, and led to improved IC_{50}s for some forms of CYP51.

A complete discussion of the approach can be found elsewhere [10]. Altogether, this example illustrates elegantly how MCSS can be used to design lead compounds de novo, especially as it represents the first example of a rationally designed ligand for any member of the cytochrome P450 family of enzymes.

Small, targeted ligand libraries

The MCSS methodology generates maps of energetically favorable positions in the binding site of a target molecule for fragments of chemical functionalities. These maps typically contain many functional group positions, allowing for a large number of potential compounds to be designed from such data. It is straightforward to make use of the MCSS methodology to construct combinatorial libraries that can efficiently explore a wide range of chemical space for the design of potential ligands, which can be screened with an efficient experimental assay. To this end, the MCSS methodology can be used to direct the construction of relatively small libraries containing molecules with a high probability of success.

In one study, the MCSS methodology was used to create a small, targeted library from functionality maps of the picornavirus capsid [9, 30]. MCSS functionality maps of the hydrophobic binding pocket of P3/Sabin poliovirus were obtained using NMA, methanol, water, acetic acid, methylammonium, magnesium ion, magnesium ion plus water, and a variety of other functional groups. The resulting MCSS minima positions compare well with compounds known to bind within the largely hydrophobic binding pocket. More importantly, the finding of new MCSS minima positions suggests that the previously known compounds did not take advantage of all available low energy conformations of its associated functional groups [21, 30].

The MCSS minima could be clustered into three distinct groups within the hydrophobic pocket, and based on this clustering, libraries of compounds were constructed that contained three individual functionalities connected by different linker moieties [30]. Of the first library, eight compounds were determined experimentally with a non-cell based assay that bound to P1/Mahoney. Four of the eight compounds were identified, and used to construct an improved library of six additional molecules, which were subsequently shown to prevent viral transformation from the native particle to its infectious intermediate using an immunoprecipitation assay. Subsequently, a crystallographic structure of one of the compounds bound to P1/Mahoney verified that the functionalities in the designed compound agreed well with the positions and orientations as calculated in the corresponding MCSS functional group maps.

Such results strongly validate the use of the MCSS methodology as a tool for structure-based drug design, and demonstrate elegantly that MCSS functionality maps can be used to design very small libraries that contain compounds that are complementary to the target molecule.

Ligand binding to nucleic acids

The three aforementioned examples considered ligand binding exclusively to proteins. However, the MCSS methodology has also been used to investigate functional group binding to nucleic acids, more specifically to RNA [40]. This is an important expansion of the application of the MCSS methodology, since the diversity of RNA tertiary structures provides the basis for specific recognition by proteins or small molecules. In the specific study illustrated here, two different RNAs, for which multiple NMR structures were available, were analyzed: (1) the TAR RNA, which contains a binding site for the amino acid arginine; and (2) the 16S rRNA, which contains a binding site for the aminoglycoside paromomycin. Following the regular MCSS methodology applied to proteins, the approach began by replicating the functional groups representing the ligands (in the case of paromomycin, only one residue was replicated), followed by a

random distribution of the replicas around the target structure and subsequent energy minimization in the force field of the target RNA.

Two sets of force field parameters were used to identify and evaluate MCSS minima. In the first set (Model 1), the non-bonded interactions between the functional group replicas and the target structure were calculated based on the CHARMM force field parameterized for nucleic acids [43], which include explicit solvation and in vacuo calculations. In a second set (Model 2), the atomic charges on the RNA phosphate groups were scaled down to mimic the presence of counterions under biological conditions. In both models, the Coulombic term of the underlying potential function and the dielectric constant of the force field were adjusted to approximate the effects of solvent screening and counterions. The search for MCSS minima was expanded from the known binding regions of the ligands to include the whole RNA, and the calculated minima were sorted based on the interaction energies, which were defined by the sum of the van der Waals (vdw) and electrostatic contributions.

Binding sites for arginine and for each residue of the aminoglycoside ligand were calculated and compared with the experimentally observed binding sites. In the case of arginine binding to the TAR RNA, the accuracy was very sensitive to the target NMR input structures. This observation is not unexpected, since differences in the NMR input structures clearly have strong effects on the MCSS calculation and therefore on the functional group mapping. In the case of paromomycin binding to the 16S rRNA, the MCSS predictions were shown to depend strongly on the geometry and electrostatic properties of the RNA binding site. Again, this does not come as a surprise, since the ligand was only partially represented (one out of four residues) in the functional group replicas. However, independently of the functional group used to represent one of the four residues of paromomycin, two RNA binding regions emerged to be more favorable. Interestingly, both regions corresponded to mono- or dicationic groups of the ligand attracted to the highly negatively charged phosphate groups within the RNA backbone.

Overall, the methodology seems to fairly accurately predict the binding of small, positively charged ligands to RNA. However, the data was stated by the authors of the study to be of preliminary nature, and to our knowledge no other example exists in the literature that utilizes the MCSS methodology to predict ligand binding to RNA. It is therefore straightforward to assume that the reliability of the predictions can surely be improved, especially to include models that give a more accurate description of the solvent and polyelectrolyte effects on the RNA-ligand interactions. Reliable predictions of known RNA binding sites will open new perspectives for the application of structure-based ligand design to these molecules, which are very attractive drug targets.

Protein flexibility and MCSS

Ligand binding, protein flexibility, and the conformational search problem

Biomolecular systems, in general, have a very large number of degrees of freedom and motion along these degrees of freedom covering a wide range of time and spatial scales [44]. These degrees of freedom and potential interactions between atoms result in a complex energy landscape representing the flexibility and conformational diversity inherent in the system [45, 46]. A biomolecular system, such as a folded protein, does not exist in a single rigid conformation, but rather assumes an ensemble of folded states [46–48].

While ligand molecules may vary in size and complexity—from simple metal ions to other macromolecules—there are many similarities between the underlying principles governing both folding and binding [45]. In fact, it has been stated that the sole difference between protein folding and ligand binding is the presence, or absence, of chain connectivity [45, 47]. In that sense, protein folding and ligand binding are very similar processes: both involve the accurate location of molecular fragments and simultaneously the exclusion of solvent and the formation of favorable hydrophobic interactions, hydrogen bonds, and electrostatic interactions [45, 49].

Binding of ligands to a specific conformation in the native state ensemble leads to relative stabilization of this conformer and a subsequent shift in the equilibrium distribution of conformers, by laws of mass action, towards a specific conformation. The extent in which ligand binding shifts the conformational distribution within the ensemble depends on the binding free energy of the interaction, and further on the differences in free energy between the individual ensemble conformations [47]. Hence designing ligands to specific members of the native state ensemble can greatly affect the overall function of the target protein by changing the distribution of conformers in the native state ensemble. A number of examples exist where ligands select pre-existing conformations from the native state ensemble of a given target protein; i.e., conformational selection [50–52].

A separate mechanism for ligand binding states that the ligand can induce the protein to adopt states that it normally would not adopt in isolation [48, 49]. The main difference between conformational selection and this “induced fit” mechanism is that conformational selection presupposes that the protein conformation that is complementary to a given binding partner exists in solution even when the

binding partner is absent. By contrast, this need not be the case with induced fit; i.e., complementary protein conformations are only formed when the protein of interest interacts with its binding partner. Recent data suggest that this latter mechanism may play an important role in some ligand binding processes, especially when one of the binding partners has regions that are disordered in solution [48, 49].

An accurate structural representation and energetic calculation of the conformational ensembles of both the free target and the unbound ligand, and how the distribution of conformational states of the target is affected in the presence of the ligand is an important, but often neglected, component of the ligand design process. The MCSS methodology provides a straightforward and general tool to address these problems in the search for novel therapeutic compounds, since protein and ligand flexibility can be incorporated into the methodology, and binding free energies can be estimated by careful evaluation of the MCSS functional group maps, as discussed earlier.

MCSS and conformational selection

Many proteins are flexible and cannot be adequately described by a single, rigid structure. If the native state of the protein is modeled as an ensemble of structures, then each of these accessible conformers of the protein corresponds to a viable target for ligand design [53]. To take advantage of the inherent flexibility, one-first develops accurate models of the native state ensemble. If a number of different structures for the protein target can be identified in the structural database, then one straightforward approach is to perform MCSS minimizations on the different X-ray crystallographic or NMR structures of the given protein target, and then to compare the functional group maps obtained from the individual structures. As different crystallographic structures may sample different side chain conformations on the target protein, integration of the functional group maps for individual structures yields a dynamic view of the protein behavior. Naturally, this approach is limited by the availability of any number of X-ray crystallographic or NMR structures, and further requires fairly intensive computational post-processing time to compare and integrate the data. However, the advantage of this approach lies in the fact that it takes into account the flexibility of the protein as a whole as derived from the input structures, and is therefore not limited to just a few flexible residues in the ligand binding site.

Alternatively, molecular dynamics (MD) simulations can be used to find different energetically favorable side chain positions in the target structure protein [54–58]. These structures can subsequently be used to calculate functionality maps for each protein conformation. In this manner several different functionality maps for a given functional group are obtained, and each map is constructed for a different target. While such approaches are potentially quite useful, it remains challenging to combine the data from functionality maps that arise from the different structures, and the whole process becomes rather time-intensive. In addition, energy minimization of only the side chains neglects any possible changes in the protein backbone, and it further remains questionable if such results will hold up unbiased upon inclusion of solvent effects.

The use of multiple protein structures that provide an ensemble of conformations, as describe earlier, seems to be the better approach to take advantage of the full flexibility of the protein. However, besides the aforementioned challenge of integrating the functional group information obtained from MCSS calculations on different conformations of the same target structure, further challenges and questions remain unanswered: What is the best source of input structures (NMR or X-ray)? How many structures are needed to include enough flexibility in the protein without undersampling? Are some structures better than others? What is the optimal number of structures to be included in the process to sample enough conformations without oversampling? How should solvent effects be included?

While the answers to these questions almost certainly will be unique for any ligand-receptor pair of interest, an illustrative example can be found for the specific case of *E. coli* ribonuclease HI [59]. While a set of 15 NMR structures and MD simulations sampled similar structures, the conformations sampled by NMR were of a wider range, showing more flexibility in both side chains and the protein backbone. When compared to two X-ray crystal structures and correlated with the available thermal factors, it was shown that multiple crystal structures covered a wider conformational range than MD simulations, but fell short of the NMR structural sampling [60–62]. In general, however, one would most likely expect this observation to be dependent on the availability and quality of the NMR and X-ray crystal structures, and further on the particular ligand-receptor pair of interest.

MCSS and induced fit

As previously discussed, induced fit assumes that the protein adopts new conformations in the presence of its binding partner—conformations that the protein would not adopt when the binding partner is not present. MCSS can be used to explore the range of conformations available to the target protein when other ligands are present. Specifically, a comparison has been made the HIV-1 protease that includes protein flexibility and compares functionality maps generated using both a rigid protein and a partially or fully flexible

protein [8]. To our knowledge, this work constitutes the single example that can be found in the literature that includes full protein flexibility in MCSS calculations to date. To investigate how different protocols for including flexibility in the target structure affect the positions of MCSS minima, the functionality maps obtained for HIV-1 protease using standard MCSS calculations with different protein structures were compared to functionality maps obtained with quenched molecular dynamics. Functionality maps using the probes methanol and methyl ammonium were obtained, since these functionalities are representative of the polar and charged chemical fragments that are utilized in standard MCSS runs.

All quenched molecular dynamics simulations that include functional group replicas were performed with the Locally Enhanced Sampling (LES) methodology [4]. As mentioned above for the case of a rigid protein, in this approximation, each functional group copy interacts with the full force field of the protein, and the protein interacts with the mean field of all of the functional group replicas, but the functional groups do not interact with each other. In case of a flexible protein, however, the minima that arise from the molecular simulations are not exact. For example, functional group minima arising from individual minimizations of the protein with one functional group replica may differ from the minima arising from quenched simulations using the LES method [6, 7]. The minima arising from these simulations are approximations to the exact result, and different metrics have been developed to determine how far removed each of these "approximate" minima are from the "true" minima which would have been obtained with repeated minimizations of a single copy of the functional group by itself [6].

In the case of HIV-1 protease, functionality maps calculated with an alternate X-ray crystal structure yielded minima with interaction energies comparable to the initial crystallographic structure [8]. However, protocols employing quenched molecular dynamics yielded minima with less favorable interaction energies relative to the minima obtained with standard MCSS and a crystallographic structure. Such a result is not unexpected as in quenched molecular dynamics simulations with the LES methodology protein atoms in the binding site see the full force field of other residues in the protein, while seeing only the mean force field of the functional group replicas. As there are several thousand replicas, the force field arising from any one copy is quite small relative to the force field arising from the protein atoms, and therefore the protein residues preferentially optimize their interactions with other protein residues, and not with the functional group replicas themselves. As a result, the functional group minima are less favored relative to the minima obtained with the original crystallographic structure.

In addition to MCSS calculations performed with an alternate crystal structure, quenched molecular dynamics simulations involving specific functional group minima within a *localized* region of the binding site on the other hand, were shown to find minima with significantly more favorable interaction energies than those found in the reference structure [8]. This suggests that locally optimized MCSS functional group minima can effectively be used to search for novel ligand conformations that further stabilize ligand binding.

Based on the example of the HIV-1 protease overall, it is suggestive that a combination of the two approaches—standard MCSS calculations with alternate structures and local optimization of MCSS minima—will be the most fruitful strategy for the design of high affinity ligands from MCSS minima. If performed consecutively, i.e., standard MCSS minimizations with different crystallographic or NMR structures of a given protein followed by local optimizations of specific minima that are of particular importance, this approach could yield highly optimized chemical fragments that could serve as the starting point for designing complete molecules that bind their targets with high affinity.

Comparison of MCSS with other fragment-based approaches

In this final section, we would like to briefly introduce two other approaches that have emerged as complementary alternatives to the MCSS methodology as tools to incorporate protein flexibility into fragment-based ligand design. A comprehensive review of these techniques is far beyond the scope of this review, but excellent overviews can be found elsewhere [1, 35, 61].

One of the earliest algorithms used to determine functional group binding sites in a target molecule is GRID [63]. As the name implies, the methodology uses a precalculated grid to represent the interaction of the ligand with the target molecule. In its initial implementation, functional groups were represented as spherical probes and the interaction energy of the probe with the target was evaluated at discrete points on a grid. Interaction energies at these grid points are evaluated using a simple potential energy function that is the sum of electrostatic, vdw, and hydrogen bond contributions [63, 64]. However, the method has evolved over the years from its initial implementation, and interaction energies for multi-atom functional groups can now be evaluated as the sum of interaction energies between single sphere probes and the protein. Based on these calculations, the program produces a contour map of the binding site of interest that can be used to identify energetically favorable fragment positions.

As such, the GRID approach is considerably different from the MCSS methodology, which represents functional groups as an atomic model for chemical fragments and evaluates the interaction energy by using the CHARMM potential energy function.

In a comparison study, MCSS and GRID functionality maps were obtained for the hydrophobic pocket of poliovirus capsid protein and the src SH3 domain using a number of different functional group probes [64]. Energetically favorable positions identified by GRID are typically found in the vicinity of MCSS minima, but MCSS finds additional minima that are not identified by low energy GRID contours in this particular study. This is not unexpected, as both approaches provide optimal positions for functional groups in the binding site, but only MCSS further provides information regarding the orientation of these functional group minima within the optimal position. In addition, GRID does not account for the intragroup energy of the functional group fragment, and therefore any internal strain that may be introduced within the fragment upon binding remains undetected. Hence, MCSS may be a more appropriate method for studying flexible chemical fragments. However, since GRID only computes interaction energies for spherical probes at pre-defined grid points in a binding site, it can compute contour maps with considerable efficiency and faster than the MCSS methodology. Given this situation, one might imagine a hybrid approach which may prove useful for obtaining functionality maps of large binding sites: to primarily obtain GRID contour maps of the binding site, and subsequently use these maps to determine the functional group positions that warrant additional, more detailed, functional group mapping with the MCSS methodology.

Another common approach used to determine complementary interactions for ligand binding makes use of the so-called pharmacophore models [65, 66]. These models aim at the three-dimensional arrangement of essential features of a particular ligand by assigning geometric rules (hydrophobic volumes and hydrogen bond vectors) and chemical properties to both ligand and receptor [67]. Typical pharmacophore features include positions on a molecule that are predominantly found to be hydrophobic, aromatic, charged, or potential hydrogen-bond donors or acceptors. These features will need to be matched with similar (i.e., hydrophobic) or opposite (i.e., charged) chemical properties in order to identify novel, complementary ligands. Often these geometric rules are based on calculations and functional group orientations derived from MCSS minimizations. As such, the pharmacophore model is a complementary tool to standard techniques, such as MCSS or docking, when the three-dimensional structure of the target has been experimentally determined. However, the technique is computationally much faster than GRID or MCSS, since the potential ligands are evaluated solely based on the condition to fit the appropriate functional group into the proper geometry pre-assigned by the pharmacophore model. Functional groups that do not fit the model are discarded. Thus, the gain in computational speed comes at the expense of more complete information of the receptor-ligand interaction, such as the inclusion of possible "induced fit" interactions at sites of the receptor that are not included in the pharmacophore model. In general, the geometrical rules that set up the pharmacophore space have to be evaluated very carefully to include all steric and electronic features necessary to ensure the optimal interactions with a specific biological target.

Conclusion

Fragment-based ligand design approaches, such as the MCSS methodology, have a number of attractive features compared to the more traditional high-throughput screening of combinatorial libraries of chemical compounds in structure-based drug design. These approaches can be applied to challenging targets, such as protein-protein interactions [11, 41], and usually result in the necessity of screening substantially fewer compounds, thereby reducing both screening times and cost drastically. Furthermore, fragment-based approaches usually detect ligands that bind specifically to a target with affinities in the μM range; such low-affinity binders are generally difficult to detect by most other methods. Moreover, fragment-based approaches result in few, if any, false positives, and non-specific binding is not an issue, since the approach detects only compounds that bind specifically to targeted fragments. Lastly, the knowledge of the preferred binding of fragments to a given target allows for lead compound optimization by growing the fragments or by combining and linking different fragments.

Most applications of MCSS have focused on proteins (Table 1), but as described earlier, the method has been applied to both DNA and RNA and fruitful results have been obtained [40, 68]. However, as the MCSS methodology only finds energetically favorable positions of *chemical fragments* in a binding site of interest, additional methods are needed to construct complete ligands based on MCSS results. To this end, a variety of different approaches, such as HOOK or Dynamic Ligand Design (DLD), have been developed to construct more complex compounds from MCSS minima as discussed elsewhere [9, 18, 69, 70]. Given the ease of use of the MCSS methodology, and the wealth of data contained in a functionality map, MCSS is a powerful tool that can be used to facilitate the design of novel therapeutic compounds.

Acknowledgments The authors would like to thank Martin Karplus who co-authored an earlier review of the MCSS methodology with C.M.S. [3]. C.M.S. is a recipient of a Burroughs Wellcome Fund Career Award in the Biomedical Sciences. C.R.S. would like to thank the Department of Chemistry at MIT for their continued research fellowship support.

References

1. Anderson AC (2003) Chem Biol 10:787. doi:10.1016/j.chembiol.2003.09.002
2. Miranker A, Karplus M (1991) Proteins 11:29. doi:10.1002/prot.340110104
3. Stultz CM, Karplus M (2006) In: Jahnke W, Erlanson DA (eds) Fragment-based approaches in drug discovery. Wiley-VCH, Weinheim
4. Elber R, Karplus M (1990) J Am Chem Soc 112:9161. doi:10.1021/ja00181a020
5. Gerber RB, Buch V, Ratner MA (1982) J Chem Phys 77:3022. doi:10.1063/1.444225
6. Stultz CM, Karplus M (1998) J Chem Phys 109:8809. doi:10.1063/1.477551
7. Ulitsky A, Elber R (1993) J Chem Phys 98:3380. doi:10.1063/1.464110
8. Stultz CM, Karplus M (1999) Proteins 37:512. doi:10.1002/(SICI)1097-0134(19991201)37:4<512::AID-PROT3>3.0.CO;2-O
9. Caflisch A, Miranker A, Karplus M (1993) J Med Chem 36:2142. doi:10.1021/jm00067a013
10. Ji H, Zhang W, Zhang M et al (2003) J Med Chem 46:474. doi:10.1021/jm020362c
11. Zeng J, Nheu T, Zorzet A et al (2001) Protein Eng 14:39. doi:10.1093/protein/14.1.39
12. Leonardi A, Barlocco D, Montesano F et al (2004) J Med Chem 47:1900. doi:10.1021/jm030944+
13. Bemis GW, Murcko MA (1996) J Med Chem 39:2887. doi:10.1021/jm9602928
14. Bemis GW, Murcko MA (1999) J Med Chem 42:5095. doi:10.1021/jm9903996
15. Brooks BR, Bruccoleri RE, Olafson BD et al (1983) J Comput Chem 4:187. doi:10.1002/jcc.540040211
16. Neria E, Fischer S, Karplus M (1996) J Chem Phys 105:1902. doi:10.1063/1.472061
17. MacKerell AD, Bashford D, Bellott M et al (1998) J Phys Chem B 102:3586. doi:10.1021/jp973084f
18. Stultz CM, Karplus M (2000) Proteins 40:258. doi:10.1002/(SICI)1097-0134(20000801)40:2<258::AID-PROT80>3.0.CO;2-I
19. Veerapandian B, Cooper JB, Sali A et al (1990) J Mol Biol 216:1017. doi:10.1016/S0022-2836(99)80017-5
20. Evensen E, Joseph-McCarthy D, Karplus M (1997) MCSSv2.1, Harvard University
21. Joseph-Mccarthy D, Hogle JM, Karplus M (1997) Proteins 29:32. doi:10.1002/(SICI)1097-0134(199709)29:1<32::AID-PROT3>3.0.CO;2-H
22. Roux B, Yu HA, Karplus M (1990) J Phys Chem 94:4683. doi:10.1021/j100374a057
23. Caflisch A (1996) J Comput Aided Mol Des 10:372. doi:10.1007/BF00124471
24. Roux B, Simonson T (1999) Biophys Chem 78:1. doi:10.1016/S0301-4622(98)00226-9
25. Stultz CM (2004) J Phys Chem B 108:16525. doi:10.1021/jp047126t
26. Baker NA (2005) Curr Opin Struct Biol 15:137. doi:10.1016/j.sbi.2005.02.001
27. Davis ME, McCammon JA (1989) J Comput Chem 10:386. doi:10.1002/jcc.540100313
28. Chothia C (1974) Nature 248:338. doi:10.1038/248338a0
29. Joseph-McCarthy D, Alvarez JC (2003) Proteins 51:189. doi:10.1002/prot.10296
30. Joseph-McCarthy D, Tsang SK, Filman DJ et al (2001) J Am Chem Soc 123:12758. doi:10.1021/ja003972f
31. Allen KN, Bellamacina CR, Ding X et al (1996) J Phys Chem 100:2605. doi:10.1021/jp952516o
32. English AC, Groom CR, Hubbard RE (2001) Protein Eng 14:47. doi:10.1093/protein/14.1.47
33. Silberstein M, Dennis S, Brown L et al (2003) J Mol Biol 332:1095. doi:10.1016/j.jmb.2003.08.019
34. Dennis S, Kortvelyesi T, Vajda S (2002) Proc Natl Acad Sci USA 99:4290. doi:10.1073/pnas.062398499
35. Kortvelyesi T, Dennis S, Silberstein M et al (2003) Proteins 51:340. doi:10.1002/prot.10287
36. Liepinsh E, Otting G (1997) Nat Biotechnol 15:264. doi:10.1038/nbt0397-264
37. English AC, Groom CR, Caves LS et al (1999) Proteins 37:628. doi:10.1002/(SICI)1097-0134(19991201)37:4<628::AID-PROT13>3.0.CO;2-G
38. Shuker SB, Hajduk PJ, Meadows RP et al (1996) Science 274:1531. doi:10.1126/science.274.5292.1531
39. Sirockin F, Sich C, Improta S et al (2002) J Am Chem Soc 124:11073. doi:10.1021/ja0265658
40. Leclerc F, Karplus M (1999) Theor Chem Acc 101:131. doi:10.1007/s002140050419
41. Zeng J, Treutlein HR (1999) Protein Eng 12:457. doi:10.1093/protein/12.6.457
42. Barbacid M (1987) Annu Rev Biochem 56:779. doi:10.1146/annurev.bi.56.070187.004023
43. Mackerell AD, Wiorkiewicz-Juczera J, Karplus M (1995) J Am Chem Soc 117:11946. doi:10.1021/ja00153a017
44. van Gunsteren WF, Bakowies D, Baron R et al (2006) Angew Chem Int Ed Engl 45:4064. doi:10.1002/anie.200502655
45. Tsai CJ, Kumar S, Ma B et al (1999) Protein Sci 8:1181. doi:10.1110/ps.8.6.1181
46. Onuchic JN, Wolynes PG (2004) Curr Opin Struct Biol 14:70. doi:10.1016/j.sbi.2004.01.009
47. Marco E, Gago F (2007) ChemMedChem 2:1388. doi:10.1002/cmdc.200700087
48. Shoemaker BA, Portman JJ, Wolynes PG (2000) Proc Natl Acad Sci USA 97:8868. doi:10.1073/pnas.160259697
49. Sugase K, Dyson HJ, Wright PE (2007) Nature 447:1021. doi:10.1038/nature05858
50. Boehr DD, Wright PE (2008) Science 320:1429. doi:10.1126/science.1158818
51. Lange OF, Lakomek NA, Fares C et al (2008) Science 320:1471. doi:10.1126/science.1157092
52. Ma B, Kumar S, Tsai CJ et al (1999) Protein Eng 12:713. doi:10.1093/protein/12.9.713
53. Teague SJ (2003) Nat Rev Drug Discov 2:527. doi:10.1038/nrd1129
54. Amaro RE, Baron R, McCammon JA (2008) J Comput Aided Mol Des 22:693. doi:10.1007/s10822-007-9159-2
55. Amaro RE, Schnaufer A, Interthal H et al (2008) Proc Natl Acad Sci USA 105:17278. doi:10.1073/pnas.0805820105
56. Cheng LS, Amaro RE, Xu D et al (2008) J Med Chem 51:3878. doi:10.1021/jm8001197
57. Durrant JD, Amaro RE, McCammon JA (2009) Chem Biol Drug Des 73:168. doi:10.1111/j.1747-0285.2008.00761.x
58. Landon MR, Amaro RE, Baron R et al (2008) Chem Biol Drug Des 71:106. doi:10.1111/j.1747-0285.2007.00614.x
59. Philippopoulos M, Lim C (1999) Proteins 36:87. doi:10.1002/(SICI)1097-0134(19990701)36:1<87::AID-PROT8>3.0.CO;2-R
60. Damm KL, Carlson HA (2007) J Am Chem Soc 129:8225. doi:10.1021/ja0709728

61. Carlson HA (2002) Curr Pharm Des 8:1571. doi:10.2174/1381612023394232
62. Clarage JB, Romo T, Andrews BK et al (1995) Proc Natl Acad Sci USA 92:3288. doi:10.1073/pnas.92.8.3288
63. Goodford PJ (1985) J Med Chem 28:849. doi:10.1021/jm00145a002
64. Bitetti-Putzer R, Joseph-McCarthy D, Hogle JM et al (2001) J Comput Aided Mol Des 15:935. doi:10.1023/A:1014309222984
65. Jacobs DJ, Rader AJ, Kuhn LA et al (2001) Proteins 44:150. doi:10.1002/prot.1081
66. Verdonk ML, Cole JC, Taylor R (1999) J Mol Biol 289:1093. doi:10.1006/jmbi.1999.2809
67. Thomas BE, Joseph-McCarthy D, Alvarez JC (2000) In: Güner OF (ed) Pharmacophore perception, development, and use in drug design. International University Line, La Jolla
68. Schechner M, Sirockin F, Stote RH et al (2004) J Med Chem 47:4373. doi:10.1021/jm0311184
69. Eisen MB, Wiley DC, Karplus M et al (1994) Proteins 19:199. doi:10.1002/prot.340190305
70. Miranker A, Karplus M (1995) Proteins 23:472. doi:10.1002/prot.340230403