Chemistry & Biology, Vol. 11, 185–194, February, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembio.2004.02.012

Systems-Based Design of Bi-Ligand Inhibitors of Oxidoreductases: Filling the Chemical Proteomic Toolbox

Daniel S. Sem,[1,2,*] Bonnie Bertolaet,[2] Brian Baker,[2,3] Edcon Chang,[2] Aurora D. Costache,[1] Stephen Coutts,[2] Qing Dong,[2] Mark Hansen,[2] Victor Hong,[2,4] Xuemei Huang,[2] Richard M. Jack,[2] Richard Kho,[2] Henk Lang,[2] Chen-Ting Ma,[2] David Meininger,[2,5] Maurizio Pellecchia,[2,6] Fabrice Pierre,[2] Hugo Villar,[2] and Lin Yu[2]

[1]Chemical Proteomics Facility at Marquette
Department of Chemistry
Marquette University
P.O. Box 1881
Milwaukee, Wisconsin 53201

[2]Triad Therapeutics, Inc.
9381 Judicial Drive
San Diego, California 92121

Summary

Genomics-driven growth in the number of enzymes of unknown function has created a need for better strategies to characterize them. Since enzyme inhibitors have traditionally served this purpose, we present here an efficient systems-based inhibitor design strategy, enabled by bioinformatic and NMR structural developments. First, we parse the oxidoreductase gene family into structural subfamilies termed pharmacofamilies, which share pharmacophore features in their cofactor binding sites. Then we identify a ligand for this site and use NMR-based binding site mapping (NMR SOLVE) to determine where to extend a combinatorial library, such that diversity elements are directed into the adjacent substrate site. The cofactor mimic is reused in the library in a manner that parallels the reuse of cofactor domains in the oxidoreductase gene family. A library designed in this manner yielded specific inhibitors for multiple oxidoreductases.

Introduction

Proteomes are inherently modular since most domains in proteins belong to superfamilies common to many organisms [1–3], and proteins are generally thought to be created by gene duplication and shuffling of a limited repertoire of domains [1, 2]. For instance, oxidoreductases frequently use the same Rossmann fold domain to bind the NAD(P)(H) cofactor but use an additional unique domain for the substrate that defines the function for a given enzyme. The presence of the conserved cofactor binding site, coupled with the modular nature of this gene family, has led us to develop a highly parallel approach to inhibitor design. In this chemical proteomic strategy, focused chemical libraries are tailored to subfamilies of large gene families to produce nM inhibitors for multiple members of the subfamily. The parallel production of inhibitors across a gene family such as the oxidoreductases will have applications in chemogenomic and functional genomic efforts to define protein functions, as well as in drug design. The latter is a significant point considering the large number of new proteins and drug targets being identified as a result of functional genomics efforts [4]. Inhibitors for various members of the oxidoreductase gene family could be used to generate chemical knockouts as a probe of protein function in vivo (chemogenomics) or of protein-ligand interactions in vitro (chemical proteomics). Such molecules, if designed for optimal ADMET (adsorption, distribution. metabolism, excretion, and toxicology) properties could even serve as early-stage leads in a parallel drug discovery program.

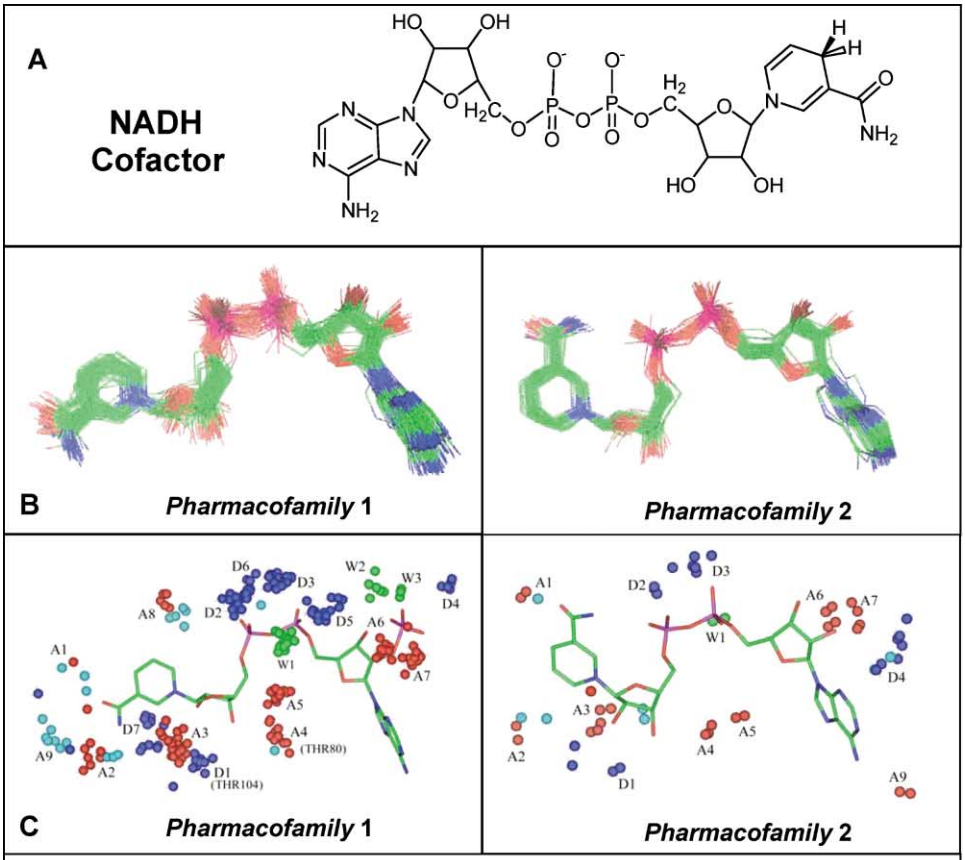
Such a systems-based approach to inhibitor design requires a solid understanding of the common binding sites in a gene family. Since cofactor conformation is a reflection of the common binding site shape, we had previously performed cluster analysis [5] on cofactors extracted from 288 oxidoreductase crystal structures. In that study [6], oxidoreductases clustered into subfamilies termed *pharmacofamilies* that were related by cofactor geometry, protein sequence, and protein fold (SCOP classification). These structural proteomic studies are being extended and used herein to enable a parallel/gene family-based approach to the design of bi-ligand inhibitors. NMR methods are used to design bi-ligand libraries off of a privileged scaffold that occupies the cofactor site, which is conserved within the *pharmacofamily*. The strategy is finally validated by the identification of potent bi-ligand inhibitors for multiple members of this *pharmacofamily*.

Results and Discussion

Structural Proteomic Analysis of Oxidoreductases: *pharmacofamilies*

The conserved cofactor geometry for oxidoreductases is apparent in Figure 1, where cofactors have been overlaid for members of the largest *pharmacofamily*, the two-domain Rossmann fold proteins. There are two subfamilies that differ only by a 180° rotation around the glycosidic bond, with *pharmacofamily 1* having the nicotinamide ring *anti* and *pharmacofamily 2* having the nicotinamide ring *syn* [6]. Especially relevant for the current studies is that this conservation of cofactor geometry is paralleled by a conservation of binding site features that describe the pharmacophore for this family. The conserved heteroatoms that define the hydrogen bond donors and acceptors that comprise this pharmacophore are shown in Figure 1C. The major oxidoreductase

*Correspondence: daniel.sem@marquette.edu
[3]Present address: Vertex Pharmaceuticals, Inc., 11010 Torreyana Road, San Diego, CA 92121.
[4]Present address, Millennium Pharmaceuticals, Inc., 40 Landsdowne Street, Cambridge, MA 02139.
[5]Present address, Tularik, Inc., 1120 Veteran's Boulevard, South San Francisco, CA 94080.
[6]Present address, The Burnham Research Institute, 10901 N. Torrey Pines Road, La Jolla, CA 92037.



D

Pharmacofamily	Enzyme Family / Example(s)	Nicotinamide Ring
1	Two-domain Rossmann fold	anti
2	Two-domain Rossmann fold	syn
3	Single-domain Rossmann fold (short-chain)	syn
4	Catalases / Monooxygenases	syn
5	β/α Barrel / Inosine monophosphate dehydrogenase	anti
6	Dihydrofolate reductases	anti
7	Disulfide dehydrogenases	anti
8	CYP oxidoreductase / ferredoxin reductase	anti
9	Isocitrate / Isopropylmalate dehydrogenase	anti
10	3-Hydroxy-3-methylglutaryl-CoA reductase	anti and syn
11	Aldehyde and shikimate dehydrogenases	anti

Figure 1. *Pharmacofamilies 1* and *2*

(A) Structure of the NADH cofactor bound by the oxidoreductases.

(B) Overlays of a unique subset of NAD(P)(H) geometries obtained from 288 crystal structures of oxidoreductases, yielding *pharmacofamilies* related by the geometry of bound cofactor. The largest families are shown here, corresponding to the two-domain Rossmann fold enzymes in *pharmacofamilies 1* (*anti*) and *2* (*syn*).

(C) Corresponding pharmacophores for *pharmacofamilies 1* and *2*, with all protein heteroatoms indicated that are within hydrogen bonding distance of atoms in the cofactor in the binding site. Regions occupied by Thr104 and Thr80 in *E. coli* DHPR (dihydrodipicolinate reductase) are indicated for reference.

(D) Summary of the major *pharmacofamilies* that were previously derived based on parsing the oxidoreductases according to geometry of bound cofactor [6]. Geometry around the C-N glycosidic bond connecting nicotinamide and ribose rings is indicated.

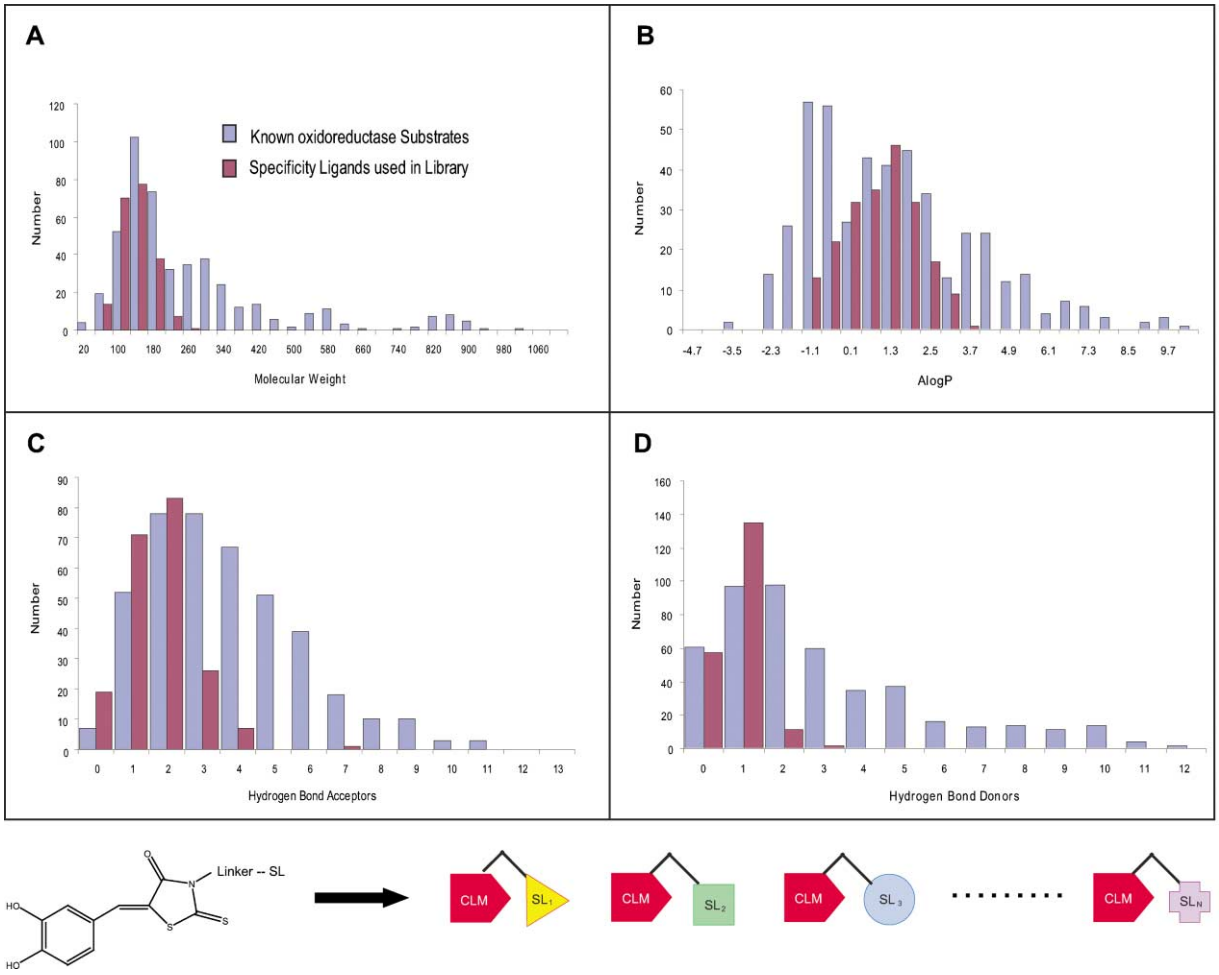


Figure 2. Range of Calculated and Predicted Physicochemical Properties of Oxidoreductase Substrates as Well as Specificity Ligands (SLs) Used in the Bi-Ligand Library

Key chemical properties of these diversity elements are compared with those for 460 known oxidoreductase substrates [7]. Calculated properties are: (A) molecular weight; (B) AlogP, a measure of hydrophobicity; (C) number of hydrogen bond acceptors; and (D) number of hydrogen bond donors.

pharmacofamilies are summarized in Figure 1D, and initial studies reported herein focus on enzymes in *pharmacofamilies 1* and *2*. The only differences between *pharmacofamilies 1* and *2* is the placement of groups around the carboxamide substituent on the nicotinamide ring. Noteworthy is the tendency of the carbonyl of the carboxamides to point in the same direction, affecting the relative placement of hydrogen bond donors (to the carboxamide C = O) and acceptors (from the carboxamide $-NH_2$) within a *pharmacofamily*.

The binding site for the nicotinamide ring of the cofactor is always close to the substrate site since the nicotinamide ring is involved in a hydride transfer reaction with the substrate. Although the binding site for the cofactor is conserved within a *pharmacofamily*, the adjacent substrate site is quite variable. This variability is reflected in the diversity of substrates acted on by oxidoreductases. We analyzed 460 oxidoreductase substrates [7] in terms of properties of interest in the drug design process [8] (Figure 2). Although most oxidoreductase substrates are in the 100–180 Da range and of modest hydrophobicity, there are a number with molecular weights in the 550 and 850 Da range, and some that are quite hydrophobic, with AlogP (Ghose and Crippen water/octanol partitioning [9]) values in excess of five.

Modular Inhibitor Design Strategy to Parallel a Modular Gene Family

Proteins that are evolutionarily related and have conserved pharmacophore features in a binding site would be expected to have similar ligand binding preferences. As such, our systems-based approach to the design of bi-ligand inhibitors of oxidoreductases begins within a *pharmacofamily*, initially chosen to be the two-domain Rossmann fold family (Figure 1) because it is the largest and most well-characterized *pharmacofamily*. Oxidoreductases, viewed in a systems-based manner, are comprised of two adjacent binding sites: the NAD(P)H cofactor (common ligand) and substrate (specificity ligand) binding sites, exemplified in Figure 3 with the enzyme dihydrodipicolinate reductase (DHPR). The inhibitor design strategy used herein parallels the modular

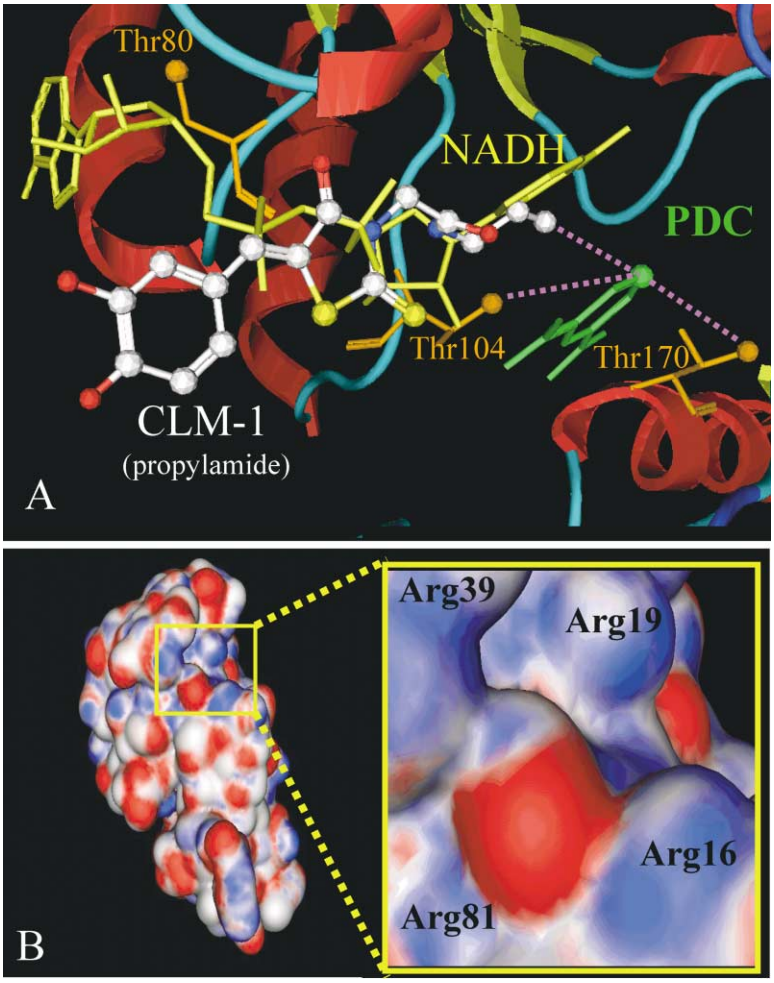


Figure 3. Comparison of Binding Modes for Computationally Docked Cofactor Analog and the NADH Cofactor

(A) Computationally docked structure of the propylamide derivative of CLM-1 (white) in the *E. coli* DHPR binding site, overlaid on the NADH structure (yellow) and adjacent to the 2,6-pyridine dicarboxylate (PDC) substrate analog (green). Docking was done with the docking algorithm contained within the MOE software package (Chemical Computing Group), with the MMFF94 forcefield and with the 1arz coordinates for DHPR [10]. Binding site threonine residues are identified in brown, with methyl groups rendered as balls. Proximity of methyl groups on the CLM's propylamide group, Thr104, and Thr170 to the PDC ligand is indicated with dashed lines.

(B) A solvent-accessible surface map color coded by partial charge (red, negative; blue, positive), with the region surrounding the negatively charged catechol oxygen shown expanded. In the expansion, red is the surface exposed para-hydroxyl group of CLM-1, and the four surrounding blue regions represent guanido groups of arginines 81, 16, 39, and 19 that approach within 5.5, 5.7, 7.0, and 9.0 Å, respectively.

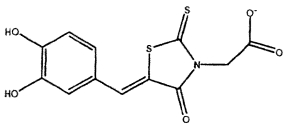
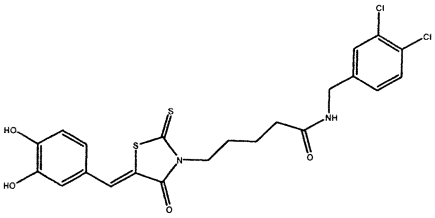
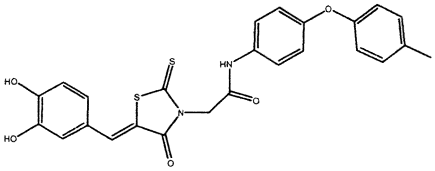
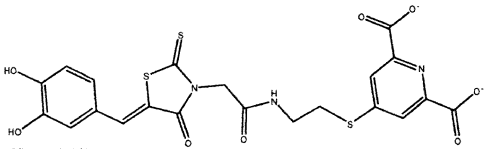
design of the oxidoreductase gene family [11] and produces inhibitors across a *pharmacofamily*, since it starts by identifying a small molecule that binds in the common ligand site (a common ligand mimic or CLM) for that class of proteins. Diversity elements are then directed from the CLM into the adjacent specificity site in the construction of a bi-ligand library.

Identification and Characterization of CLMs

We selected CLM candidates computationally by matching the pharmacophore properties of the nicotinamide mononucleotide portion of NADH bound to DHPR [10], an oxidoreductase in *pharmacofamily 1*, and an enzyme essential for cell wall synthesis in *Mycobacterium tuberculosis* [12]. This ligand-based search employed the icosahedral matching algorithm [13] contained within the THREEDOM software package (Interprobe, Inc.) to identify potential inhibitors, which were then purchased and tested against DHPR as well as other dehydrogenases in this *pharmacofamily*. The most drug-like and crossreactive of these were resynthesized and retested. Crossreactivity is a desired property, since a CLM is effectively a privileged scaffold that is going to provide baseline affinity across a *pharmacofamily*, with further increases in affinity later achieved for specific targets by directing bi-ligand libraries into the variable substrate pocket. One of the inhibitors identified in this screening process was modified to produce a more potent and soluble analog by replacing a phenyl ring with an acetic acid group, resulting in CLM-1 (Table 1; Figure 4). The modeled structure of a propylamide derivative of CLM-1 is shown docked into the binding site of DHPR and overlaid on NADH (Figure 3A). The docked structure binds in a mode that differs from that originally predicted based on direct comparisons to cofactor, which may be a reflection of an inherent symmetry in the NADH molecule that has a nicotinamide ring on one end and an adenine ring on the other. Indeed, another low-energy docked structure had the propylamide group in the adenine site, but the orientation shown here with the propylamide group proximal to the substrate site is most consistent with the NMR SOLVE data described below. The electrostatic surface shown in Figure 3B indicates that the catechol ring is somewhat solvent exposed and surrounded by positive charge density from adjacent arginines.

All computationally selected CLM candidates were commercially available and tested for binding potency through steady-state kinetic inhibition studies with DHPR, with a representative set of compounds shown in Figure 4A. CLMs that bind at the NAD(P)(H) site were identified based on inhibition profiles. For example,

Table 1. Affinity and Specificity of CLM and Bi-Ligand Molecules for Oxidoreductases in *pharmacofamilies 1* and *2*

Structure[a]	LDH[b]	DHPR[b]	DOXPR[b]
	55 μM	26 μM	>50 μM
	**42 nM**[c]	>50 μM	10 μM
	12 μM	>25 μM	**202 nM**[c]
	620 nM	**100 nM**[c]	7.9 μM

[a] While the SL of the first bi-ligand was condensed with the carboxylic acid of the linker in Figure 5C, the other 2 SLs were condensed with the acid of the shorter linker on CLM-1 [15].
[b] Numbers are K_{is} values except for DOXPR, which has an IC_{50}, which should approximate a K_{is}. LDH (lactate dehydrogenase) and DHPR are in *pharmacofamily 1*, while DOXPR (1-deoxy-D-xylulose-5-phosphate reductoisomerase) is in *pharmacofamily 2* [6].
[c] Most potent inhibition value amongst the three enzymes is indicated in bold.

CLM-1 is a competitive inhibitor versus NADPH and, therefore, likely binds in the cofactor binding site. Interestingly, the inhibition pattern showed a squared dependence on concentration, suggesting that some synergy might exist between sequential binding events to the DHPR tetramer. The fit was best to a competitive model, with no apparent intercept effect in double reciprocal plots (Figure 4B).

Determining the Expansion Point for the Bi-Ligand Library with NMR SOLVE

We then experimentally determined the CLM's orientation and relative position in the cofactor binding site by the NMR SOLVE method [11, 14]. The NMR SOLVE method begins by mapping a binding site relative to a reference ligand, such as cofactor, and then characterizing the binding mode of a novel ligand, such as a CLM, relative to the reference ligand. Key information obtained with NMR SOLVE is where a linker should be placed such that chemical diversity elements can be attached and directed into an adjacent specificity pocket. In order to avoid problems with spectral overlap in 2D NMR experiments, studies were performed on deuterated and sparsely labeled protein, with the ^{1}H-^{13}C label present only in the methyl groups of the 8-Met, 16-Ile, and 14-Thr residues in DHPR. Previously, we had mapped the DHPR binding site with NADH as a reference ligand [11] (Figure 5A). Now, we used these assignments for key binding site residues to orient a CLM candidate relative to where the reference cofactor had bound. Although the acid version of the best CLM (CLM-1, Table 1 and Figure 4) showed interaction primarily between its catechol ring and the distal Thr80 (data not shown), the propylamide derivative in Figure 5B showed a clear nuclear Overhauser effect (NOE) between the terminal methyl of the CLM and the 2,6-pyridinedicarboxylate (PDC) substrate antagonist, which itself showed an NOE to the Thr104 interface residue. PDC also showed an NOE to a residue assigned to the substrate site (Thr170). PDC is a stable analog of the reactive dihydrodipicolinate substrate and is used here to help mimic the ternary complex that would normally form in the steady-state catalytic cycle. Based on the NOE to PDC, the end of the propylamide functionality therefore appears to be the appropriate place for attaching a specificity ligand. We then used a carboxylic acid functionality here as a bi-ligand library expansion point. To verify that this functionality also resides in the substrate site, we compared 2D HMQC spectra (Figure 5C) for each version of this CLM, with and without the terminal carboxylic acid (red and blue crosspeaks). This chemical perturbation of the ligand caused the largest changes in the crosspeak for Thr170, a residue known to reside in the substrate site, and to a residue at the interface of the cofactor and substrate sites (Thr104). These data largely confirmed the docked structure in

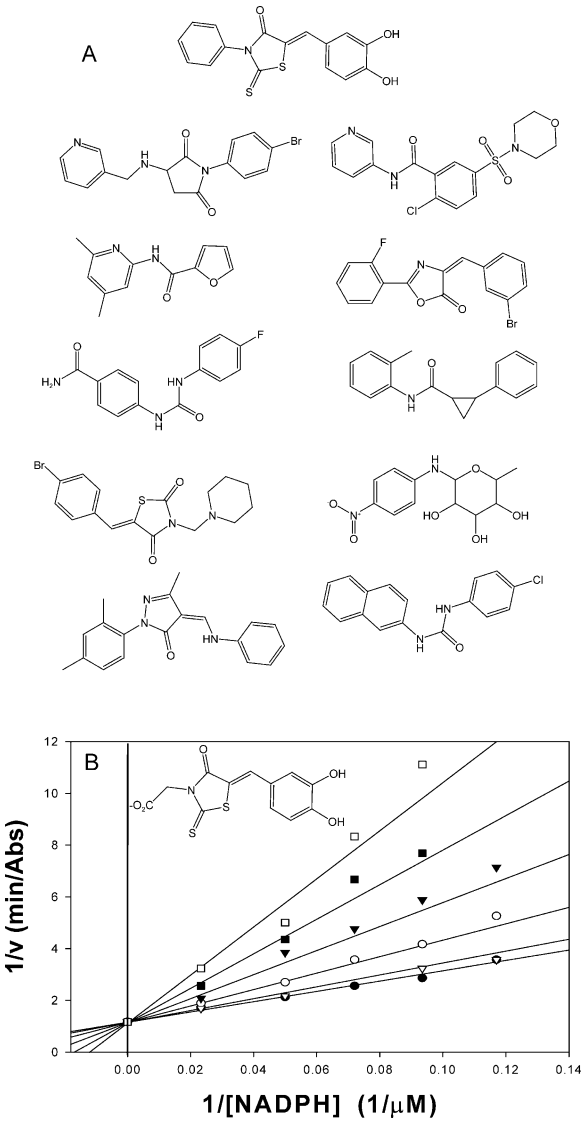


Figure 4. Computationally Selected Cofactor Analogs and Steady-State Characterization as a Cofactor Analog

(A) A representative set of 11 computationally selected and tested cofactor mimics, with the top structure being the only one that showed significant inhibitory activity.

(B) Steady-state inhibition profile for CLM-1, which is a modified form of the top inhibitor in (A). Profile is with inhibitor versus NADPH. Profile represents the fit to the equation for a competitive inhibitor with a squared dependence on inhibitor concentration. The fit gave a K_{is} value of 26 ± 2 μM. Curves for alcohol dehydrogenase (K_{is} = 101 μM), lactate dehydrogenase (K_{is} = 55 μM), and DOXPR (K_{is} > 50 μM) fit best to a model for competitive inhibition.

Figure 3A, since the terminal methyl of the propylamide is that part of the CLM that is closest to PDC (within 5 Å). It should be noted that the NMR SOLVE experiments would have suggested the same library expansion point in the absence of any protein structural information, since crosspeaks for threonine residues were assigned based on proximity to reference ligands (NADH and PDC). That is, it was never necessary to assign Thr104 to a specific residue number, as it would have been adequate to view it only as the crosspeak for the residue at the interface of the NADH and PDC binding sites. Structural data were included here only to illustrate the method. The NMR data not only suggested the bi-ligand library expansion point, but it also confirmed that PDC binding mode was not significantly altered in the CLM:PDC:DHPR ternary complex (compared to the NADH:PDC:DHPR ternary complex), since PDC showed the same pattern of NOEs to threonine methyl protons in both complexes. Finally, the selective perturbation of active site methyl proton chemical shifts in the complex with CLM and the complex with the bi-ligand (see below) allow us to rule out any nonspecific mechanism for inhibition that could have produced competitive inhibition profiles, at least for these inhibitors.

Validation of the NMR-Selected Library Expansion Point

Based on these NMR data, CLM-1 was linked to PDC, the specificity ligand analog. The corresponding bi-ligand compound had a K_{is} of 100 nM, which represents a 250-fold increase in affinity over the starting CLM (Table 1). The common and specificity sites were both occupied, based on NMR chemical shift mapping studies showing perturbations of residues in both binding sites (Figure 5D). Although the squared effect on inhibition complicated steady-state analysis, analogous bi-ligands were made with the same specificity ligand but with variants of the CLM that gave less pronounced cooperativity effects. Steady-state kinetic profiles for two of these bi-ligand molecules are shown in Figure 6, showing a best fit to a competitive inhibition model versus both the NADPH and dihydrodipicolinate substrates, as expected for a bi-ligand inhibitor.

Building the Focused Bi-Ligand Library

The last step of this systems-based bi-ligand design process involves the addition of diversity elements to the CLM such that they are directed into the substrate (specificity ligand [SL]) site in the manner suggested by NMR SOLVE. Although the geometric relationship of the CLM and SL sites is conserved in oxidoreductases, the actual size and electrostatic properties of the SL binding site will vary in a way that parallels the diversity of substrates used by oxidoreductases, as discussed above. The diversity elements attached to the conserved CLM scaffold were matched to the properties of substrates used throughout the oxidoreductase gene family in that they roughly paralleled the distribution of the chemical properties surveyed in Figure 2.

We selected 300 diversity elements (commercially available) and chemically joined them to the CLM-linker construct. The resulting bi-ligand library was then screened against three Rossmann-fold enzymes in *pharmacofamilies 1* and *2*: DHPR, lactate dehydrogenase (LDH), and 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DOXPR). The starting CLM bound only weakly to the three enzymes, with K_{is} values in the 25–100 μM range (Table 1). However, after adding the diversity elements in the position selected with NMR SOLVE, steady-state enzyme kinetic screening identified a specific bi-ligand inhibitor of LDH with a K_{is} of 42 nM and a best fit to a model for competitive inhibition. This represents

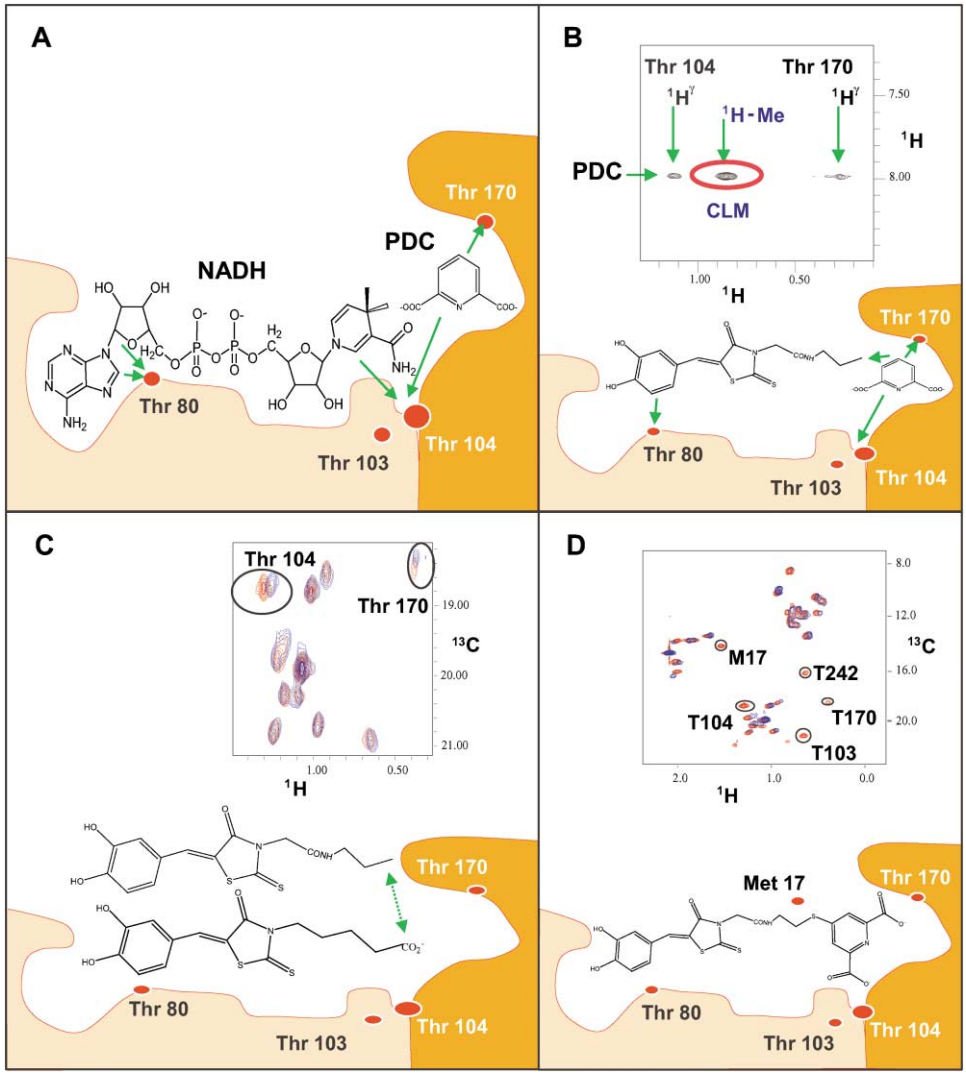


Figure 5. NMR SOLVE Data for DHPR

(A) The binding site is mapped relative to the NADH cofactor, identifying NMR probe atoms [14].

(B) NOESY data for a CLM-1 analog, which places the methyl terminus of the propylamide functionality closest to the SL site.

(C) Chemical alteration of the end of this alkyl chain in creating CLM-2 produces changes to crosspeaks for atoms in the SL site in the overlay of HMQC spectra for complexes of DHPR with both versions of the CLM (red and blue).

(D) HMQC spectra of DHPR in the absence (red) and presence (blue) of the bi-ligand inhibitor.

an increase in affinity over the starting CLM of a thousand-fold, with most of the increased affinity directed toward specificity interactions, since the bi-ligand binds two to three orders of magnitude stronger to LDH than to DHPR or DOXPR. We also identified a bi-ligand inhibitor that bound with an IC_{50} of 202 nM to DOXPR, also with selectivity. Based on these data, we propose that a bi-ligand collection of sufficient size and diversity, built with an appropriately chosen CLM and well placed linkers, will produce nM inhibitors for most members of a *pharmacofamily*.

Thermodynamic Foundation for Gene Family Focused Bi-Ligand Libraries

Fragment-based assembly strategies are a very efficient means of designing inhibitors [16–18]. The systems-based bi-ligand design process described herein makes fragment assembly more likely to produce inhibitors for multiple, related proteins, since the CLM fragment provides a baseline of affinity across a *pharmacofamily* (ΔG_{CLM}). Furthermore, since the mere joining of two molecular fragments has been proposed to provide as much as 45 entropy units, corresponding to an increase in affinity of 10^8-fold [19, 20] associated with the "chelate effect," the combined effect of adding a CLM to an SL could be as large as $10^8/K_{CLM}$ fold. This could only occur for those SLs binding in the specificity pocket adjacent to the CLM, thus ensuring that linkage with the CLM provides specificity for a given oxidoreductase. Although the full magnitude of the chelate effect is still being investigated by researchers, it is in any case quite large and reports of enhancements in affinity or rates

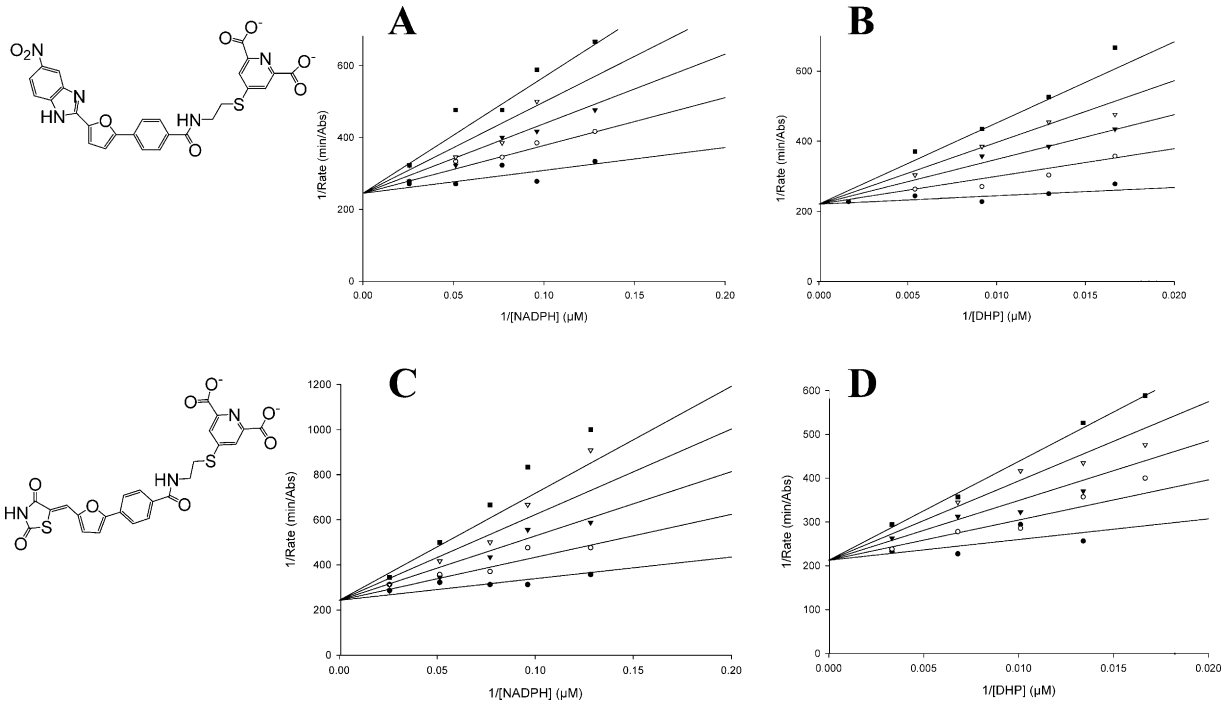


Figure 6. Steady-State Inhibition Profiles for Bi-Ligand Molecules with CLMs that Are Variants of CLM-1

Bi-ligand structures are shown to the left of their respective inhibition profiles. Curves fit best to equations where inhibition was competitive versus both cofactor (NADPH) and substrate. Enzyme was *E. coli* DHPR. The first bi-ligand was varied (0, 400, 750, 1100, 1500 nM) versus NADPH and gave a K_{is} of 370 ± 90 nM (A), and versus dihydrodipicolinate (DHP) gave a K_{is} of 170 ± 50 nM (B). The second bi-ligand was varied (0, 500, 1000, 1500, 2000 nM) versus NADPH and gave a K_{is} of 500 ± 100 nM (C), and versus dihydrodipicolinate gave a K_{is} of 530 ± 140 nM (D).

approaching 10^8-fold have been reported [21]. In practice, only a tiny fraction of this affinity increase will ever be fully realized because linkers might be flexible, produce nonoptimal placement of the CLM and SL ligands, or have repulsive interactions with a binding site. Still, this combinatorial strategy for fragment linkage is an efficient way of focusing a library, since even imperfect linkage can produce large affinity boosts for multiple enzymes within a *pharmacofamily*.

Significance

Genomes have now been sequenced for numerous organisms, and expectations are that this information will provide a better understanding of biology, as well as yield new and better therapeutics. This cannot be realized until more efficient strategies are developed to characterize proteins, and tools are created to translate this knowledge into inhibitors as mechanistic probes and drugs. To this end, we had previously reported a structural proteomic analysis of oxidoreductases, which comprise 2%–4% of most proteomes [6]. Herein, we presented a systems-based strategy for designing inhibitors of oxidoreductases. This strategy parallels nature's modular approach for designing the oxidoreductase gene family itself. Chemical libraries designed in this manner could be used as a source of chemogenomic probes for defining functions of members of the oxidoreductase gene family, as well as a source of drug leads for the many new targets being identified in functional genomics efforts.

Experimental Procedures

Computational Search for CLMs

Computational methods were used to identify CLM candidates to be tested as cofactor mimics. Coordinates of the nicotinamide mononucleotide (NMN) portion of the NADH cofactor were extracted from the structure of cofactor in complex with DHPR (pdb code: 1arz). These coordinates were used to search against databases of commercially available compounds (such as from ASINEX [Moscow, Russia]), which contained molecules with precalculated three-dimensional structures. The search was quite fast since it involved only matching the shape of the NMN portion of the cofactor to the precalculated structures in the small molecule database through an icosahedral matching algorithm [13, 22]. Still, an ellipsoid shape prefilter was used as a first screen to identify molecules that were roughly the shape of NMN to prescreen and eliminate obviously poor matches. To address heteroatom composition and hydrogen bonding capabilities at an approximate level, hybrid shape-matching scores were used. Shapes were compared both as a function of all atoms and as a function of heteratoms only, with scores weighted equally in the hybrid score. After prefiltering the circa 40,000 molecules, shape-matching scores were calculated with the THREEDOM software package by comparing full structures or only the heteratoms. These two sets of scores were combined to create an average "hybrid" score through Perl scripts developed in-house. Generally around 5% of the computationally selected compounds were found to inhibit one of the three dehydrogenases tested. Although searches against the full-length NADH cofactor produced a somewhat higher hit rate (up to 10% versus lactate dehydrogenase), these compounds were not used due to their higher molecular weight and unsuitability as drug leads.

Computational Docking of the Propylamide Derivative of CLM-1

The propylamide derivative of CLM-1 was docked into the crystal structure of the DHPR/NADH/2,6-PDC ternary complex after removing the cofactor. First, the ligand was minimized by AM1 (Gaussian98) with a net charge of –1, localized on the para-hydroxy group of the catechol ring. Docking was then performed with the MOE software package (Chemical Computing Group) and the MMFF94 forcefield, with ligand flexible and protein kept rigid. Protein was also energy minimized (MMFF94) before docking, and then 25 docking runs were performed with random starting orientations. Optimization was with simulated annealing, with an initial temperature of 1000 K and six cycles per run. Docking calculations included protein atoms within a 62 Å × 62 Å × 62 Å box surrounding the site previously occupied by cofactor. The choice to dock into the cofactor site was based on the observation of competitive inhibition patterns (with the more soluble CLM-1) versus cofactor in the steady-state enzyme kinetic studies described below. One of the two lowest energy structures is shown in Figure 3A, with the NADH structure overlaid back in its original orientation so that the relative binding mode of CLM-1 derivative and "reference" cofactor can be compared. Although the lowest energy structure had the propylamide group extending into the adenine site, only the orientation shown in Figure 3A was consistent with the NMR SOLVE data described below. A solvent-accessible surface map color coded by partial charge is shown in Figure 3B, with the region surrounding the negative (red) catechol ring shown in the expansion. The four surrounding blue regions represent guanido groups of arginines 81, 16, 39, and 19 that approach within 5.5, 5.7, 7.0, and 9.0 Å, respectively, of the charged oxygen of the catechol ring.

Synthesis of Bi-Ligand Library

The 3,4-dihydroxyphenylmethylene-rhodanine CLM (CLM-1) was synthesized by heating a solution of 13.6 g 3,4-dihydroxybenzaldehyde, and 19.3 g rhodanine acetic acid at 95°C in 200 ml acetic acid for 6 hr. After cooling the solution, the precipitate was collected and washed with acetic acid (2 × 5 mL) to give 20 g of product CLM-1. The CLM-1 acid (or amine; 1.5 eq) and PS-carbodiimide resin (2 eq) were reacted in THF (10 mL/g) for 1 hr. The desired amine (or acid) to be conjugated to the CLM-1 was then added (1 eq) and reacted overnight. Product was extracted twice with THF and solvent evaporated to give desired product.

Protein Production

The *E. coli* DHPR protein was expressed with a pET21a vector (Novagen) and purified as described previously [14]. Briefly, DHPR was uniformly enriched in ^{2}H and ^{15}N and contained $^{1}H/^{15}N/^{13}C$-labeled threonine, $^{1}H/^{13}C(\delta$-methyl)-labeled isoleucine, and $^{1}H/^{13}C$ (ϵ-methyl)-labeled methionine, and was produced through a modified version of supplemented M9 minimal media. Purification was on a Q Sepharose anion exchange column (Amersham). The DOXPR gene was cloned from *E. coli* genomic DNA by PCR utilizing the following primers: 5′-GCCACTGCATATGAAGCAACTCACCATTCTGG and 3′-GCCACTGGGATCCTCAGCTTGCGAGACGCATC. DOXPR was expressed and purified as reported [14, 23].

Steady-State Inhibition Studies

All reactions were monitored spectrophotometrically at 340 nm by using initial rates from the first 5% of reaction. Absorbance changes at 340 nm are from production or consumption of NAD(P)H cofactor. LDH reaction mixtures contained 100 mM Hepes buffer (pH 7.4), 2.5 mM pyruvate, 10 μM NADH, and 5 ng/mL lactate dehydrogenase. DOXPR reaction mixtures contained 100 mM Hepes buffer (pH 7.4), 1.2 mM DOXP, 8 μM NADPH, 1 mM $MnCl_2$, and 10 μg/mL DOXPR. DHPR reactions were as described previously [24], with either NADH or NADPH cofactor. In all cases when screening bi-ligands, nonvaried substrate concentrations were kept close to their K_m values. Data were fitted to appropriate equations for competitive, noncompetitive, and uncompetitive inhibition through nonlinear least-squares fitting [25]. Mechanism of inhibition was established through analysis of the overall sigma value for the fit, as well as the magnitude of standard deviations for inhibition constants (K_{ii}, K_{is}). The inhibition pattern for CLM-1 (Table 1) showed a squared dependence on inhibitor concentration (Figure 4B), still fitting best to a competitive model:

$$v = \frac{V_{max}\,A}{A + K_m\,(1 + (I/K_{is})^2)}$$

where v is the initial velocity, I is inhibitor concentration, K_{is} is the slope inhibition constant, A is cofactor (NADH) concentration, V_{max} is the maximum velocity, and K_m is the Michaelis constant.

NMR Spectroscopy

NMR experiments were performed on a Bruker DRX700 spectrometer equipped with a triple resonance probe and triple axis gradient coil. Tetrameric DHPR concentration was ~75 μM (300 μM monomer) in 25 mM Tris-D_{11} buffer (pH 7.8) at a temperature of 303 K, with a sample volume of 150 μl in Shigemi tubes, as described previously [14]. Most NMR experiments were performed in the presence of the PDC substrate analog along with either NADH or CLM, in order to mimic the active ternary complex that is produced in the steady-state catalytic cycle. PDC was not used for studies of bi-ligands or for the inhibitors being compared in Figure 5C. PDC is a fairly potent substrate analog, with a K_{is} value of 26 μM versus the dihydrodipicolinate substrate [24]. Selective WURST adiabatic decoupling [26] of the γ from the β ^{13}C of Thr was used in NMR experiments to decrease overlap of the 14 Thr residues. Typical 2D [^{13}C,^{1}H] HMQC spectra were recorded in 30 min. Typical 2D [^{1}H,^{1}H] NOESY spectra were acquired with 256 × 2048 complex points and with mixing times between 50 and 500 ms. ^{13}C decoupling during acquisition was with a GARP composite decoupling sequence [27], while $^{13}C\gamma$ decoupling during the evolution period was with a 180° refocusing pulse. Ambiguities due to proton overlap among Thr and Met methyl proton chemical shifts were removed by recording a 3D [^{13}C,^{1}H] resolved [^{1}H,^{1}H] NOESY experiment [28]. NOEs were later verified as not being due to spin diffusion through the QUIET-NOESY experiment [29, 30].

Acknowledgments

We thank Dr. John Blanchard (Albert Einstein College of Medicine, NY) for the original pET11a DHPR expression construct. The work described in this article was performed at Triad Therapeutics, with the exception of the modeling in Figure 3, which was done by A.D.C. at Marquette University.

Received: September 10, 2003
Revised: November 17, 2003
Accepted: November 17, 2003
Published: February 20, 2004

References

1. Bashton, M., and Chothia, C. (2002). The geometry of domain combination in proteins. J. Mol. Biol. *315*, 927–939.
2. Apic, G., Gough, J., and Teichmann, S.A. (2001). An insight into domain combinations. Bioinformatics *17*, S83–S89.
3. Todd, A.E., Orengo, C.A., and Thornton, J.M. (2001). Evolution of function in protein superfamilies, from a structural perspective. J. Mol. Biol. *307*, 1113–1143.
4. Drews, J. (1998). In Search of Tomorrow's Medicines (New York: Springer-Verlag).
5. Willett, P. (1987). Similarity and Clustering in Chemical Information Systems (Letchworth, UK: Research Studies Press).
6. Kho, R., Baker, B.L., Newman, J.V., Jack, R.M., Sem, D.S., Villar, H.O., and Hansen, M.R. (2003). A path from primary protein sequence to ligand recognition. Proteins *50*, 589–599.
7. You, K.S. (1985). Stereospecificity for nicotinamide nucleotides in enzymatic and chemical hydride transfer reactions. CRC Crit. Rev. Biochem. *17*, 313–451.
8. Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. *46*, 3–26.
9. Ghose, A.K., and Crippen, G.M. (1987). Atomic physicochemical

parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. J. Chem. Inf. Comput. Sci. *27*, 21–35.
10. Scapin, G., Reddy, S.G., Zheng, R., and Blanchard, J.S. (1997). Three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase in complex with NADH and the inhibitor 2,6-pyridine dicarboxylate. Biochemistry *36*, 15081–15088.
11. Sem, D.S., Yu, L., Coutts, S.M., and Jack, R. (2001). Object-oriented approach to drug design enabled by NMR SOLVE: First real-time structural tool for characterizing protein-ligand interactions. J. Cell. Biochem. *S37*, 99–105.
12. Cirillo, J.D., Weisbrod, T.R., Banerjee, A., Bloom, B.R., and Jacobs, W.R., Jr. (1994). Genetic determination of the meso-diaminopimelate biosynthetic pathway of *mycobacteria*. J. Bacteriol. *176*, 4424–4429.
13. Bladon, P. (1989). A rapid method for comparing and matching the spherical parameter surfaces of molecules and other irregular objects. J. Mol. Graph. *7*, 130–137.
14. Pellecchia, M., Meininger, D., Dong, Q., Chang, E., Jack, R., and Sem, D.S. (2002). NMR-based structural characterization of large protein-ligand interactions. J. Biomol. NMR *22*, 165–173.
15. Parlow, J.J., Mischke, D.A., and Woodard, S.S. (1997). Utility of complementary molecular reactivity and molecular recognition (CMR/R) technology and polymer-supported reagents in the solution-phase synthesis of heterocyclic carboxamides. J. Org. Chem. *62*, 5908.
16. Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovery of high affinity ligands for proteins: SAR by NMR. Science *274*, 1531–1534.
17. Fejzo, J., Lepre, C.A., Peng, J.W., Bemis, G.W., Ajay, Murcko, M.A., and Moore, J.M. (1999). The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. Chem. Biol. *6*, 755–769.
18. Erlanson, D.A., Braisted, A.C., Raphael, D.R., Randal, M., Stroud, R.M., Gordon, E.M., and Wells, J.A. (2000). Site-directed ligand discovery. Proc. Natl. Acad. Sci. USA *97*, 9367–9372.
19. Page, M.I., and Jencks, W.P. (1971). Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. Proc. Natl. Acad. Sci. USA *68*, 1678–1683.
20. Snider, M.J., Lazarevic, D., and Wolfenden, R. (2002). Catalysis by entropic effects: the action of cytidine deaminase on 5,6-dihydrocytidine. Biochemistry *41*, 3925–3930.
21. Carlow, D., and Wolfenden, R. (1998). Substrate connectivity effects in the transition state for cytidine deaminase. Biochemistry *37*, 11873–11878.
22. Doucet, J.-P., and Weber, J. (1996). Computer-Aided Molecular Design: Theory and Applications (San Diego: Academic Press).
23. Meininger, D.P., Rance, M., Starovasnik, M.A., Fairbrother, W.J., and Skelton, N.J. (2000). Characterization of the binding interface between the E-domain of staphylococcal protein A and an antibody Fv-fragment. Biochemistry *39*, 26–36.
24. Reddy, S.G., Sacchettini, J.C., and Blanchard, J.S. (1995). Expression, purification, and characterization of *Escherichia coli* dihydrodipicolinate dehydrogenase. Biochemistry *34*, 3492–3501.
25. Cleland, W.W. (1979). Statistical analysis of enzyme kinetic data. Methods Enzymol. *63*, 103–138.
26. Kupce, E., and Wagner, G. (1995). Wideband homonuclear decoupling in protein spectra. J. Magn. Reson. *B109*, 329–333.
27. Shaka, A.J., Barker, P.B., and Freeman, R. (1985). Computer-optimized decoupling scheme for wideband applications and low-level operation. J. Magn. Reson. *64*, 547–552.
28. Cavanagh, H.J., Fairbrother, W.J., Palmer, A.G., III, and Skelton, N.J. (1996). Protein NMR Spectroscopy, Principles and Practice, First Edition (New York: Academic Press).
29. Neuhaus, D., and Williamson, M.P. (2000). The Nuclear Overhauser Effect in Structural and Conformational Analysis, First Edition (New York: Wiley-VCH).
30. Vincent, S.J.F., Zwahlen, C., and Bodenhausen, G. (1996). Suppression of spin diffusion in selected frequency bands of nuclear Overhauser spectra. J. Biomol. NMR *7*, 169–172.