PROTEINS: Structure, Function, and Bioinformatics 64:60–67 (2006)

Inhibition of Protein–Protein Interactions: The Discovery of Druglike β-Catenin Inhibitors by Combining Virtual and Biophysical Screening

Jean-Yves Trosset,[1]* Claudio Dalvit,[1] Stefan Knapp,[1] Marina Fasolini,[1] Marina Veronesi,[1] Sergio Mantegani,[1] Laura M. Gianellini,[2] Cornel Catana,[1] Michael Sundström,[1] Pieter F.W. Stouten,[1] and Jürgen K. Moll[2]
[1]*Department of Chemistry, Nerviano Medical Sciences, Nerviano, MI, Italy*
[2]*Department of Biology, Nerviano Medical Sciences, Nerviano, MI, Italy*

ABSTRACT **The interaction between β-catenin and Tcf family members is crucial for the Wnt signal transduction pathway, which is commonly mutated in cancer. This interaction extends over a very large surface area (4800 Å^2), and inhibiting such interactions using low molecular weight inhibitors is a challenge. However, protein surfaces frequently contain "hot spots," small patches that are the main mediators of binding affinity. By making tight interactions with a hot spot, a small molecule can compete with a protein. The Tcf3/Tcf4-binding surface on β-catenin contains a well-defined hot spot around residues K435 and R469. A 17,700 compounds subset of the Pharmacia corporate collection was docked to this hot spot with the QXP program; 22 of the best scoring compounds were put into a biophysical (NMR and ITC) screening funnel, where specific binding to β-catenin, competition with Tcf4 and finally binding constants were determined. This process led to the discovery of three druglike, low molecular weight Tcf4-competitive compounds with the tightest binder having a K_D of 450 nM. Our approach can be used in several situations (e.g., when selecting compounds from external collections, when no biochemical functional assay is available, or when no HTS is envisioned), and it may be generally applicable to the identification of inhibitors of protein–protein interactions. Proteins 2006; 64:60–67.** © 2006 Wiley-Liss, Inc.

Key words: β-catenin; Tcf4; docking; virtual screening; druglike inhibitors; inhibition of protein–protein interaction

INTRODUCTION

Proteins are involved in interactions with an estimated average of 5–10 protein partners.[1] The protein–protein interface exhibits an overall match between polar residues through a large hydrogen bond network, often mediated by water molecules, and hydrophobic interactions between aromatic and aliphatic patches.[2] This poses great challenges for disrupting large protein interfaces with a small molecule.

However, possibilities exist to exploit the presence of small and narrow pockets to make strong interactions. Wells and coworkers showed, in fact, that most of the protein–protein-binding enthalpy is due to interactions with a small number of so-called "hot spots," well-characterized patches on the surface of the proteins.[3,4] Engineered miniproteins with a small number of relevant residues, which are essential for binding, have confirmed this concept.[5] The first attempts to inhibit protein–protein interactions involved mimicking one of the partners with a short peptide.[6] This was followed by the discovery of small molecular weight inhibitors of protein complexes such as LFA1/ICAM,[7,8] IL-2/IL-2 receptor,[9,10] and P53/MDM2[11] (for reviews, see Berg,[12] Ocke and Gadek,[13] and Norin and Sundström[14] and references therein).

Disruption of the interaction between β-catenin and Tcf/LEF members is considered to be a valid anticancer strategy because the pathway is constitutively switched on in several cancer types including colorectal cancer. The complex of β-catenin with Tcf3 or Tcf4 is characterized by an unusually large binding interface[15,16] with a binding constant of K_D ~10 nM. Disrupting such a tight binding complex with a small organic molecule is challenging, but it may be possible by targeting one or more hot spots.

β-catenin is a superhelix with a large and shallow positively charged groove. The N-terminal 56 residues of Tcf4 or the close analogue Tcf3 bind along this groove with a total binding surface of about 4800 Å^2. The N-terminal strand of the β-hairpin observed in Tcf3/β-catenin crystal structure[15] (Fig. 1) is disordered in the Tcf4/β-catenin complex.[16] Key residues of Tcf4 have been identified by studying the binding of β-catenin with various derivatives of Tcf4 including single-Ala mutants.[18–20] Three residues emerged from these studies as major contributors to the binding energy: L48 (which is part of the C-terminal helix of Tcf4 and interacts with the N-terminal part of β-catenin; see corresponding Tcf3 helix in green at the bottom of

S. Knapp's and M. Sundström's present address is Structural Genomics Consortium, Botnar Research Centre, University of Oxford, Oxford OX3 7LD, UK.

C. Catana's present address is CADD, Ann Arbor Laboratories, Pfizer Global R&D, 2800 Plymouth Road, Ann Arbor, MI 48105.

*Correspondence to: Jean-Yves Trosset, Department of Chemistry, Nerviano Medical Sciences, Viale Pasteur 10, 20014, Nerviano, MI, Italy. E-mail: jytrosset@gmail.com

Received 24 June 2005; Revised 18 November 2005; Accepted 28 November 2005

Published online 27 March 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20955

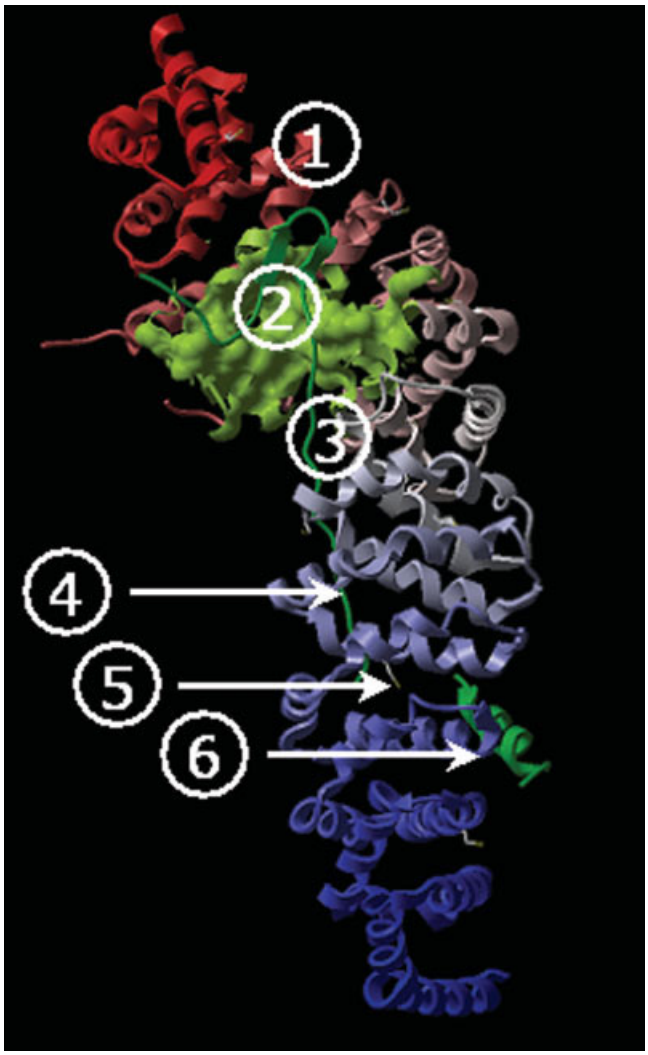

Fig. 1. β-catenin/Tcf3 crystal structure.[15] The encircled numbers refer to the six β-catenin cavities described in the text (Materials and Methods, Hot Spot Selection). Tcf3 is shown as a green ribbon and β-catenin as a ribbon whose color ranges from blue (N-terminus) to red (C-terminus). The hot spot area near K435/R469 (cavity 2) is shown as a solid green surface. The figure was generated with Molsoft's ICM.[17]

Fig.1), D16, and D11. The D11 residue is not observed in the electron density maps of the β-catenin/Tcf4 complex,[16] whereas its corresponding analog E11 of Tcf3, located on the tip of a β-hairpin in the β-catenin/Tcf3 complex, is engaged in solvent-exposed interactions with polar residues on β-catenin. The critical residue D16 is well resolved in both the β-catenin/Tcf3 and β-catenin/Tcf4 complex structures and well conserved among all β-catenin binding partners. This negatively charged amino acid makes specific, partially buried interactions with K435, one of the two key charged residues of β-catenin involved in interactions with other proteins (the other residue is K312). This area of β-catenin is less shallow than the rest of the molecular surface and is most suitable as a hot spot for structure-guided drug design.

Drug discovery projects with a known target structure can be initiated by virtual screening in various ways. External compound collections can be virtually screened, and only compounds that are predicted to inhibit the target will be acquired. Virtual libraries can be screened in silico, and the results can be used to select scaffolds and to help design the final library to be synthesized. In addition, when no high-throughput-screening (HTS) is envisioned or has been carried out yet, a discovery effort can be jump-started by virtually screening corporate and/or public compound collections. When no biochemical or other functional assay is available, the combination of virtual screening and biophysical screening can be an alternative, and in this work we pursued that approach to identify putative β-catenin inhibitors. The virtual screening experiment consisted of docking a subset of the Pharmacia collection to the K435-R469 hot spot region, estimating the binding energy of all compounds ("scoring"), and visually inspecting the most promising candidates. The surviving in silico hits were subsequently assessed for binding to and competition with Tcf4 using NMR WaterLOGSY and isothermal calorimetry (ITC) experiments.

MATERIALS AND METHODS

Hot Spot Selection

Identification of hot spots on the surface of β-catenin was initially performed with the PASS[21] program using the crystal structure of the β-catenin/Tcf3 complex.[15] Identical results were produced with the β-catenin/Tcf4[16] complex after its crystal structure had become available. Six distinct cavities can be located in the β-catenin/Tcf3 interface (Fig. 1). Cavity 1 is small and narrow and is located near G575, R582, and R612 of β-catenin. It interacts with the top of the Tcf3 β-hairpin (E11). This area does not contain nearby pockets able to anchor a small compound. Moreover, this region of β-catenin does not interact with Tcf4 according to recent crystallographic studies of the β-catenin/Tcf4 complex. Therefore, we excluded this hotspot as a potential binding site. Cavity 2 is the largest one and is part of the groove formed by a kink between armadillo repeats 9 and 10 of β-catenin. On both sides, there are two small pockets, one containing K435, which interacts with D16 of Tcf4, and the second pocket located near E571 and K508. These two pockets are about 12 Å apart and separated by a flat area located above the guanidine group of R469. Cavity 3 is formed by R386, P463, and T418, and it interacts with Tcf3-F21. Because cavity 3 is more solvent exposed than cavity 2, this cavity is a secondary hot spot candidate. Cavity 4 contains K312 and binds to Tcf3-E24 (Tcf4-E29). It is much more solvent exposed, smaller, and more shallow than cavities 2 and 3. Cavities 5 and 6, identified by PASS, are even more shallow and are located near L41 and L48 of the N-terminal α-helix of Tcf3. To select the hot spot for the virtual screening, a more detailed analysis of the six cavities was carried out by using the solvent-accessible surface calculated by the FLO_QXP[22] package. The K435/R469 region of β-catenin contains several small clefts that can potentially anchor a small molecular weight inhibitor and was, therefore, selected as a hot spot for docking.

We also carried out HyPARE[23] calculations on the Tcf4/β-catenin complex. HyPARE relates the association rate constant k_{on} to the magnitude of electrostatic forces using the Debye-Hückel theory for electrolyte solutions.[24] The calculations indicate that Tcf4-D16 makes the largest contribution to the Tcf4/β-catenin association rate constant. This is consistent with affinities determined for binding of β-catenin to various Tcf4 mutants.[18–20] Because Tcf4-D16 interacts with β-catenin-K435, these results provide further support for the selection of the region near K435 as the hot spot to target.

Initial Compound Selection

The following filters were applied to approximately 90,000 compounds from the Pharmacia & Upjohn collection: 1) no violations of Lipinski's Rules of Five[25] for 96% of the compounds; 2) inventory $>$ 5 mg; 3) no structural features that make a compound unsuitable or undesirable for screening; 4) low biological promiscuity (active in $<$4 non-antibiotic screens); and 5) solubility $>$ 50 μM. For the solubility filter, either experimentally determined (turbidimetric or nephelometric) values or in silico solubility predictions with proprietary Pharmacia models[26,27] were used. This led to a subset of 17,700 compounds that entered the virtual screening.

Virtual Screening

The FLO_QXP package was used to flexibly dock the 17,700-compound set to our chosen hot spot. The docking region included all residues within a distance of 12 Å from the δC of R469 using the atomic coordinates of the β-catenin/Tcf3 complex.[15] Side-chains of C429, D390, C466, and K508 of β-catenin were allowed to move during the docking. Each compound was docked by using 300 steps of MCDock. This docking algorithm uses repeated cycles of Monte Carlo perturbation of ligand geometry (rotatable dihedral angles and Cartesian coordinates of flexible ring atoms) followed by energy minimization to generate an ensemble of docked ligands. The procedure does not optimize just one ligand geometry but a diverse ensemble of low-energy poses (i.e., the RMS atomic distance between any pair of conformers $>$ 0.5 Å). The best 10 poses for each compound were stored for further analysis. Poses that exceeded empirical limits for the internal ligand energy with respect to the global minimum of the free ligand in vacuum (15 kJ/mol) and the Van der Waals repulsion between ligand and receptor (17 kJ/mol) were discarded. Poses with a total QXP association energy (Eass) −30 kJ/mol and contact energy (Ecnt) −24 kJ/mol were also discarded. The surviving poses were rank-ordered according to Eass and separately according to Ecnt. Visual inspection of the top ranking 3000 compounds led to the selection of compounds that show strong interactions in more than one part of the receptor's binding region.

Protein Purification

The armadillo repeat region (residues 134–671) in β-catenin, cloned as a BamH1/Not-1 fragment into the polylinker of pGex-6p2 (Stratagene) in frame with the Gst-precision protease, was expressed in *Escherichia coli* (BL21). Cells were grown to midlog phase at 37°C, and protein expression was induced by addition of 0.1 mM IPTG. Induced cells were then grown for 4 h and harvested by centrifugation. Cell pellets were resuspended in lysis buffer (PBS, 2 mM DTT), lysed by a French press, and purified by glutathione sepharose affinity chromatography. The Gst-fusion protein was eluted by using 20 mM reduced glutathione dissolved in phosphate-buffered saline (PBS). The purified proteins were dialyzed into precision cleavage buffer (50 mM Tris/Cl pH 7.6, 1 mM EDTA, 150 mM NaCl) and subsequently cleaved overnight by using 10 U of precision protease per milligram of recombinant protein. The Gst tag, uncleaved protein, and the protease were removed by a second step of glutathione sepharose affinity chromatography. The protein was 95% pure as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Tcf4 (1-53 W-His_6) was cloned into the NdeI and Xho1 sites of pET21b. An additional tryptophan residue was introduced between the His_6 tag and the protein to facilitate the determination of the protein concentration by UV spectroscopy. The additional C-terminal residues did not alter the binding affinity of Tcf4 for β-catenin/armadillo. The protein was purified by following the protocol provided by Novagen.

NMR Screening

WaterLOGSY is a method for the identification of compounds that interact with macromolecules.[28,29] In this experiment, the large bulk water magnetization is partially transferred via the protein-ligand complex to the free ligand in a selective manner. Binding compounds are easily identified in the spectra by looking at the sign of their resonances. Compounds that do not bind show negative or no peaks at all, whereas binders show positive peaks. The compounds were screened in mixtures consisting of four or five molecules. Mixtures that gave a positive signal were deconvoluted to identify individual binders. The β-catenin armadillo repeat concentration used in the NMR experiments was 2 μM in 5 mM Tris, 10 mM NaCl pH 7.3. The Tcf4 concentration was 25 μM. Compounds were dissolved in deuterated dimethyl sulfoxide (DMSO) at a concentration of 40 mM. They were first screened at 20°C in mixtures at 50 μM concentration. The samples for the competition experiments were buffered in PBS. NMR experiments were performed with a Varian Inova 600 MHz spectrometer equipped with a 5-mm triple-resonance probe and a Sample Management System (SMS) autosampler. For each sample a reference spectrum and a WaterLOGSY spectrum was recorded.

ITC Measurements

ITC measurements were carried out at 20°C by using a VP-ITC titration calorimeter (MicroCal). Samples were extensively dialyzed against PBS, 1 mM DTT. Each titration experiment consisted of an initial 2-μL injection, followed by 20 injections of 5 μL. Heats of dilution were measured in blank titrations by injecting Tcf4 (1-53 W-

His_6) into the buffer used for each particular experiment. These were subtracted from the binding heats. Direct binding experiments were carried out at 3 μM of β-catenin and 60 μM for the compound. Concentrations for competition assays were 100 μM for Tcf4 (1-53 W-His_6) and 5 μM for β-catenin. Each titration experiment was performed on mixtures of four compounds, where each compound was present in a concentration of 100 μM. Binding heats were determined by a nonlinear least-squares fit of the experimental measurements using the Origin software package (MicroCal) and assuming a single binding site model. Compound mixtures that showed at least a threefold reduction in Tcf4 binding affinity were selected for further characterization.

RESULTS AND DISCUSSION

Virtual Screening

The docking with QXP and subsequent energy filtering and visual inspection of binding modes led to a final selection of 42 candidates out of the original 17,700 compounds. As indicated above, several selection criteria were used: association energy, contact energy, intramolecular ligand strain, Van der Waals repulsion, and visual inspection. We had found previously that true binders occur among compounds with favorable association energy, favorable contact energy, or both, so it is important to select molecules on the basis of either criteria. Past experience also showed that compounds with too unfavorable values for two individual QXP energy terms (ligand strain and Van der Waals repulsion) are unlikely to be active, irrespective of their calculated overall binding affinity. For that reason, we placed hard limits on these two terms: poses with ligand strain > 15 kJ/mol or Van der Waals repulsion > 17 kJ/mol were eliminated. Although the energies we calculate should suffice to make an optimal selection, we and others (e.g., Wang et al.[30]) have found repeatedly that visual inspection is essential for the selection of active compounds. For example, using the solvent-accessible (SA) atomic surface generated by FLO_QXP proved useful for the identification of tight binders, whereas overall energies or individual energy terms were less discriminating. SA atomic surfaces are very powerful in visual inspection for two reasons. 1) If the center of a ligand heavy atom lies on an SA surface, it makes an optimal interaction. If one displays a Van der Waals surface instead, it is very difficult to assess whether a ligand atom makes an optimal interaction. 2) Although an SA molecular surface is smoother than an SA atomic surface and, therefore, easier to inspect visually, we found that it conceals details that are critical for the identification of optimal interactions. Visual inspection also entailed the identification and preferential selection of compounds with multiple anchor points (i.e., several tight contacts made by different parts of the molecule), which we thought could be important when targeting large protein interface with small ligands. There are probably also other binding features that an expert modeler intuitively recognizes during visual inspection, but which apparently are difficult to quantify and not adequately captured by common scoring functions.

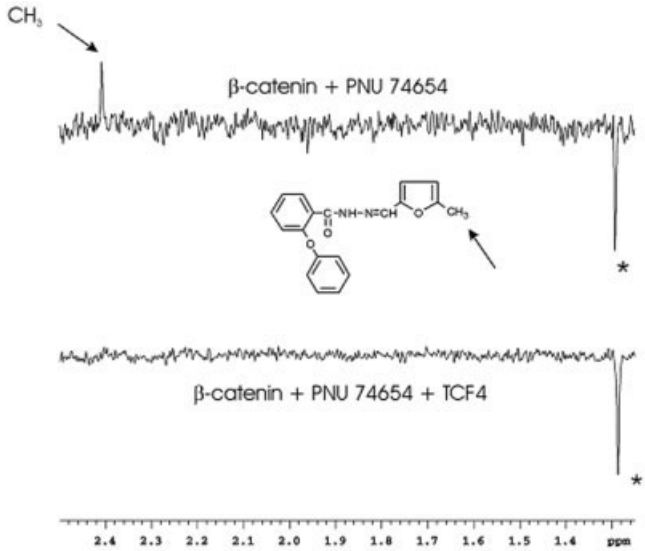

Fig. 2. Expanded region of the WaterLOGSY spectra recorded for β-catenin and PNU-74654 in the absence (top spectrum) and in the presence (bottom spectrum) of Tcf4. The arrow indicates the PNU-74654 methyl group signal. With Tcf4 present in the solution, this resonance is missing (bottom spectrum). The NMR spectrum was recorded with 800 (top spectrum) and 2048 scans (bottom spectrum). A large number of scans were acquired for the solution in the presence of Tcf4 to detect the complete displacement of PNU-74654 from β-catenin.

Biophysical Screening

Of the 42 selected compounds, 22 were available in sufficient quantity to be submitted to biophysical studies for binding and competition measurement using NMR WaterLOGSY and ITC. NMR WaterLOGSY identified 7 compounds as β-catenin binders. Independently and contemporaneously, the 22 compounds were screened in Tcf4 competition assays using ITC. Two compound mixtures resulted in an about 10-fold reduction of Tcf4-binding affinity, whereas the other compound mixtures had no measurable effect. The two active compound mixtures contained three of the seven hits detected by the NMR screening effort. Because it is not expected that compounds that bind weakly (>10 μM) to the Tcf4 interaction site show measurable inhibition of Tcf4 binding, the remaining four hits found by NMR are either only interacting weakly with β-catenin or they do not compete with Tcf4. Overall, of 22 tested compounds, 3 were confirmed as binders and Tcf4 competitors, corresponding to a hit rate of 14%.

Biophysical and Biological Characterization of PNU-74654[31,32]

The most active hit from the NMR and ITC screen is PNU-74654. Its structure and part of its WaterLOGSY competition spectrum are shown in Figure 2. The positive methyl group resonance (2.4 ppm) of the compound (indicated by an arrow) shows that PNU-74654 is a binder to the target. The resonance peak with opposite sign at 1.3

ppm (labeled with an asterisk) is due to a marker, which does not interact with the protein. Competition was observed by adding 25 μM Tcf4 to the solution. This resulted in the disappearance of the signal at 2.4 ppm (Fig. 2, bottom), which is consistent with complete displacement of PNU-74654 from β-catenin by Tcf4 and confirms the ITC competition results (10-fold reduction of Tcf4 affinity for β-catenin). Control WaterLOGSY experiments performed for PNU-74654 in the absence of β-catenin resulted in no effect for the signal at 2.4 ppm (data not shown). Disappearance or attenuation of the methyl group signal of PNU-74654 by other compounds can potentially be used to identify new inhibitors.

Direct binding ITC experiments were carried out to determine a binding constant for PNU-74654. Dilution heats were determined by a blank titration and were similar to heat effects measured after saturation of the binding site. A binding constant of $2.2 \pm 0.9\ 10^6\ M^{-1}$, ($K_D$ = 450 nM) with an associated enthalpy change ΔH of -2.0 ± 0.5 kcal/mol and a stoichiometry of 1:1 were derived from the binding isotherm (Fig. 3).

Cellular activity using a Luciferase reporter system for Tcf-4 transactivation showed specific inhibition in the presence of PNU-74654 (data not shown).

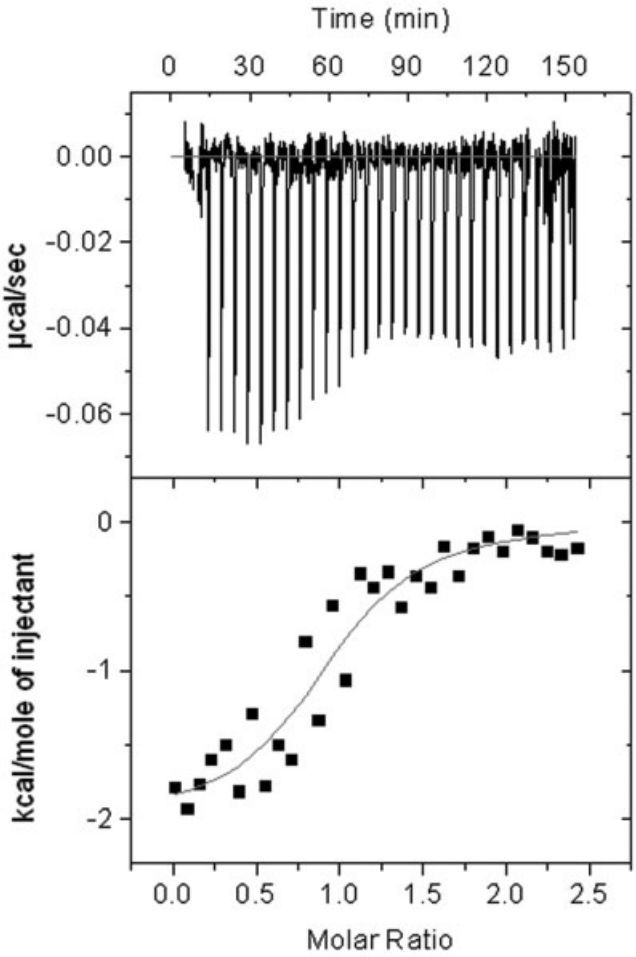

Fig. 3. Experimental calorimetric data of the binding of PNU-74654 to the armadillo repeat region of β-catenin. The top pane shows the raw heat data obtained over a series of injections of PNU-74654 to a 5 μM β-catenin/armadillo solution. The integrated heat signals of the data from the top pane gave rise to the binding curve shown in the bottom pane. The solid line represents a calculated curve using the best-fit parameters obtained by a nonlinear least-squares fit, after having subtracted the separately determined heats of dilution from the raw binding heats.

Binding Mode Determination of PNU-74654 and Hot Spot Confirmation

To determine the binding mode of PNU-74654 to β-catenin, various crystallization experiments (soaking and co-crystallization) were initiated, but they were unsuccessful. Using published crystallization conditions,[15,16] crystals of β-catenin without ligand were produced. Therefore, we refined our binding mode hypothesis using results from extensive docking studies and experimental binding data on the lead compound PNU-74654 and two of its analogs. These analogs were identified through a similarity search of our corporate collection; in analog 1, the methyl group on the furane (see arrow in Fig. 2) has been replaced by a proton, and in analog 2, the distal phenyl group has been replaced by a piperidine moiety. Analog 1 exhibits a decrease of affinity (data not shown), and analog 2 does not exhibit any competition with Tcf 4. These extensive docking experiments were carried out with FLO_QXP 2003, whose new FullDock+ and MCDock+ modules have better algorithms and scoring functions than FLO_QXP 1999, which was used to perform the original virtual screening.

The predicted binding mode of the lead compound PNU-74654 in our selected hot spot is shown in Figure 4. One of the pockets (left yellow ellipse in Fig. 4) contains the methylfurane, whereas the other (right ellipse) contains the phenyl moiety. Experimental data for PNU-74654 and its two analogs show that modifying either the methylfurane or phenyl group causes a clear decrease in binding affinity, providing evidence that both groups are strongly interacting with the protein. The predicted affinity of the pose in Figure 4 was decomposed into atomic contributions with the FLO_QXP profile program. The methyl group on the furane is fully buried in a narrow cleft and indeed makes a contribution to the affinity that is significantly higher than the contribution from any other PNU-74654 atom. All binding modes that stem from the extensive docking studies involve binding to one or both pockets. Poses with interactions with only one of those pockets have either the furane or phenyl group exposed to solvent and fail to explain the decreased experimental affinities of both analogs. Therefore, we concluded that it is most likely that the binding mode of PNU-74654 involves the two narrow pockets on either side of this hot spot. The pose of Figure 4 has the best predicted binding affinity. However, the pose where the methylfurane and phenyl groups trade places (involving a 180° rotation of the entire molecule) has a binding affinity that is only slightly worse.

We cannot exclude the possibility that PNU-74654 binds to other potential hot spots, notably cavity 3 around R386. However, unlike the β-catenin hot spot around K435/R469 (cavity 2), the R386 hot spot does not contain any clefts to which a small molecule can anchor itself well. In addition,

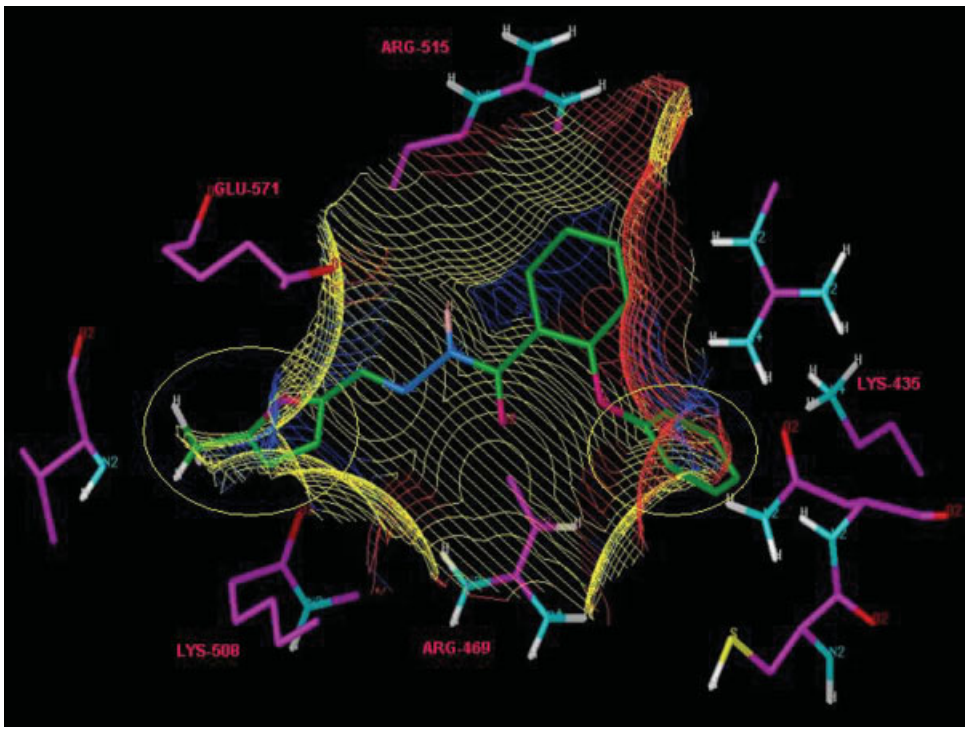

Fig. 4. Proposed binding mode of PNU-74654 (green) to the K435/R469 hot spot of β-catenin. The two narrow cavities are marked by yellow ellipses. The one on the right interacts normally with D16 of Tcf4. Right- and left-hand cavities correspond to the groove of α-helices 23 and 30 of the armadillo repeat, respectively. The solvent-accessible surface has been calculated with the QXP surface program. Color coding is as follows: yellow = hydrophobic; red = hydrogen bond donor; and blue = hydrogen bond acceptor.

binding to that hot spot cannot explain the observed decrease in binding affinity for the two PNU-74654 analogs. And finally, QXP scores derived from extensive docking experiments suggest that PNU-74654 binds more than 1000 times tighter to cavity 2 than to cavity 3 (but docking energies often overestimate true binding affinities).

Recently, Lepourcelet et al.[33] biochemically screened a chemical library and identified eight antagonists of the β-catenin/Tcf4 complex with in vitro IC_{50} values in the low μM range. To elucidate their binding modes, five (CGP049090, PKF115-584, PKF118-744, PKF222-815, ZTM000990; see Fig. 2, p. 91 of Lepourcelet et al.[33]) were subjected to extensive docking experiments to three hot spots around K435/R469 (cavity 2), R386 (cavity 3), and K312 (cavity 4) by using 1000 Monte Carlo steps of MCDock+. To accommodate CGP049090, PKF115-584, and PKF222-815 (which are significantly larger than PNU-74654), exposed side-chains in and around the cavities were allowed to move, but no global rearrangements (i.e., movement of entire secondary structure elements) were permitted. According to the QXP contact energy, all compounds fit best to the primary hot spot near K435/R469, and they bind at least 160 times tighter to that hot spot than to the R386 or K312 hot spots. Although the larger Lepourcelet compounds do not enter the two adjacent pockets of the K435/R469 hot spot (cavity 2) as deeply as PNU-74654, their large central ring systems make more extensive hydrophobic contact with the floor of the hot spot, which may explain their potency.

The docking results indicate that all compounds we studied in detail bind best to the K435/R469 hot spot. The NMR and ITC results show that PNU-74654 competes with Tcf4, and mutagenesis had already shown that the K435/R469 hot spot contributes significantly to Tcf4 binding. The binding affinities of two PNU-74654 analogues are consistent with the binding mode we propose. Combining all these pieces of evidence make a compelling argument for the selection of the hot spot and the correctness of the binding modes we propose.

Assessing the Performance of the Virtual Screen

In a typical docking experiment, we would use QXP with the MCDock+ algorithm and carry out 300–500 Monte Carlo cycles per compound. On a single 3.4 GHz Intel EM64T processor, this takes about 30 s. On a typical 56-processor cluster, one can, therefore, dock more than 1 million compounds per week (fewer if one considers multiple tautomers and protonation states per compound). Other labor-intensive and time-consuming parts of the process are the pre- and post-processing phases. The former phase encompasses analysis of the protein, setting it up for docking, and preparing the small molecules, which will take around 1–3 days, provided one is already familiar with the structural features of the target protein. The post-processing phase, which includes numerical filter-

ing and analysis of the results by visual inspection, may take another 2 days. Thus, overall virtual screening is fast and inexpensive.

The hit rate of 14% we obtained is higher than what would typically be obtained with experimental HTS, but that comparison may be misleading. The approach described here contains a filtering step (virtual screening) and a confirmation step (NMR/ITC), and the hit rate refers only to the latter step. The hit rate of experimental HTS typically refers to the entire process (the initial screen plus the confirmation). Ultimately, hit rates are not the most critical parameters. More important are the absolute numbers of confirmed active compounds or compound classes that a process (be it virtual or experimental) yields and the overall time and costs associated with that process.

Another key aspect is the novelty and druglikeness of the compounds that are found. In this context, it is noteworthy that our combination of virtual and biophysical screening has identified potent, low molecular weight β-catenin inhibitors, while we know of no druglike compounds that have been identified by HTS. Because we had filtered our compounds prior to docking and 96% of the compounds in our virtual screening collection do not violate any Lipinski rule, we were all but guaranteed to find druglike molecules, but that does not explain why no druglike inhibitors were found by HTS. In our opinion, experimental HTS and virtual screening are complementary in that active compounds are often identified by one and not the other, so ideally these techniques are used in parallel.

CONCLUSIONS

The combination of virtual screening and medium-throughput biophysical (NMR and ITC) assays proved to be a powerful approach for the identification of compounds that disrupt the interaction between Tcf4 and β-catenin, a putative oncology target. Using various methods, we determined and confirmed the most promising hot spot, the β-catenin site near K435/R469, which contributes most to the interaction with Tcf4. Subsequently, a collection of compounds was docked to this hot spot with the QXP docking program. After applying several energy criteria and visual inspection, 22 compounds with adequate inventory were assayed. Three of these bind to β-catenin and compete with Tcf4, representing a 14% hit rate. The tightest binder (PNU-74654) has a K_D of 450 nM. Further docking studies with PNU-74654 and binding data on related compounds supported the binding modes we proposed and also revealed that compounds that make interactions with more than one pocket tend to bind tighter. We have demonstrated that small molecules can interfere with protein–protein interactions, that such small molecules can be found by virtual screening, and that they can rapidly be confirmed biophysically. It is of interest that other publications had indicated that identifying small, druglike β-catenin inhibitors by HTS is difficult. The combined approach we describe here may well be turned into a generally applicable strategy for the identification of inhibitors of protein–protein interactions, which may make a large number of therapeutic targets accessible for intervention.

ACKNOWLEDGMENTS

We thank Mic Lajiness and John Van Drie for preparing the subset of our compound collection, Arndt Schnuchel for providing recombinant Tcf4 (1-53 W-His_6), and Francesco Colotta for critically reading the manuscript. We thank all of them also for helpful and stimulating discussions.

REFERENCES

1. Drews J. Drug discovery: a historical perpective. Science 2000;287: 1960–1964.
2. Jones S, Thornton J. Principles of protein–protein interactions. Proc Natl Acad Sci USA 1996;93:13–20.
3. Arkin ME, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. Nat Rev Drug Discov Sci 2004;3:301–317.
4. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. Science 1995;267:383–386.
5. Cochan G. Protein-protein interfaces: mimics and inhibitors. Curr Opin Chem Biol 2001;5:654–659.
6. Peczuh MW, Hamilton AD. Peptide and protein recognition by designed molecules. Chem Rev 2000;100:2479–2494.
7. Kallen J, Welzenbach K, Ramage P, Geyl D, Kriwacki K, Legge G, Cottens S, Weitz-Schmidt G, Hommel U. Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain. J Mol Biol 1999;292:1–9.
8. Gadek TR, Burdick DJ, McDowell RS, Stanley MS, Marsters JC Jr, Paris KJ, Oare DA, Reynolds ME, Ladner C, Zioncheck KA, Lee WP, Gribling P, Dennis MS, Skelton NJ, Tumas DB, Clark KR, Keating SM, Beresini MH, Tilley JW, Presta LG, Bodary SC. Generation of an LFA-1 antagonist by the transfer of the ICAM-1 immunoregulatory epitope to a small molecule. Science 2002 295:1086–1089
9. Emerson SD, Palermo R, Liu CM, Tilley JW, Chen L, Danho W, Madison VS, Greeley DN, Ju G, Fry DC. NMR characterization of interleukin-2 in complexes with the IL-2Rα receptor component and with low molecular weight compounds that inhibit the IL-2/IL-Rα interaction. Protein Sci 2003;12:811–822.
10. Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, Raphael DR, Taylor L, Wang J, McDowell RS, Wells JA, Braisted AC. Binding of small molecules to an adaptive protein-protein interface. Proc Natl Acad Sci USA 2003;100:1603–1608.
11. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 2004;303:844–848.
12. Berg T. Modulation of protein-protein interaction with small organic molecules. Angew Chem. 2003;42:2462–2481.
13. Ocke DA, Gadek TR. Inhibitors of protein-protein interactions. Expert Opin Ther Patents 2002;12:393–400.
14. Norin M, Sundström M. Protein models in drug discovery. Curr Opin Drug Discov Dev 2001;4:284–290.
15. Graham TA, Weaver C, Mao F, Kimelman D, Xu W. Crystal structure of a β-catenin/Tcf complex. Cell 2000;103:885–896.
16. Poy F, Lepourcelet M, Shivdasani RA, Eck MJ. Structure of a human Tcf4-β-catenin complex. Nat Struct Biol 2001;12:1053–1057.
17. Abagyan R, Totrov M, Kuznetsov D. ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem 1994;15:488–506.
18. Omer CA, Miller PJ, Diehl RE, Kral AM. Identification of Tcf4 residues involved in high-affinity β-catenin binding. Biochem Biophys Res Comm 1999;256:584–590.
19. Knapp S, Zamai M, Volpi D, Nardese V, Avanzi N, Breton J, Plyte S, Flocco M, Marconi M, Isacchi A, Caiolfa VR. Thermodynamics of the high-affinity interaction of Tcf4 with β-catenin. J Mol Biol 2001;306:1179–1189.
20. Fasolini M, Wu X, Flocco M, Trosset JY, Oppermann U, Stefan K. Hotspots in Tcf4 for the interaction with β-catenin. J Biol Chem 2003;278:21092–21098.

21. Brady GP Jr, Stouten PFW. Fast prediction and visualization of protein binding pockets with PASS. J Comput Aided Mol Des 2000;14:383–401.
22. McMartin C, Bohacek RS. QXP: powerful, rapid computer algorithms for structure-based drug design. J Comput Aided Mol Des 1997;11:333–344. QXP_FLO program: Thistlesoft, PO Box 227, Colebrook, CT 06021.
23. Shaul Y, Schreiber G. Exploring the charge space of protein-protein association: a proteomic study. Proteins 2005;60:341–352. http://bip.weizmann.ac.il/hypareb/main.
24. Schreiber G, Fersht AR. Rapid, electrostatically assisted association of proteins. Nat Struct Biol 1996;3:427–431.
25. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 2001;46:3–26.
26. Gao H, Shanmugasundaram V, Lee P. Estimation of aqueous solubility of organic compounds with QSPR approach. Pharm Res 2002;19:497–503.
27. Catana C, Gao H, Orrenius C, Stouten PFW. Linear and non-linear methods in modeling the aqueous solubility of organic compounds. J Chem Inf Model 2005;45:170–176.
28. Dalvit C, Fogliatto GP, Stewart A, Veronesi M, Stockman B. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. J Biomol NMR 2001;21:349–359.
29. Dalvit C, Fasolini M, Flocco M, Knapp S, Pevarello P, Veronesi M. NMR-based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands. J Med Chem 2002;45:2610–2614.
30. Wang JL, Liu D, Zhang ZJ, Shan S, Han X, Srinivasula SM, Croce CM, Alnemri ES, Huang Z. Structure-based discovery of an organic compound that binds Bcl-2 protein and induces apoptosis of tumor cells. Proc Natl Acad Sci USA 2000;97:7124–7129.
31. Rector DL, Conder GA, Folz SD. Preparation of N-heterocyclylalkylidene acylhydrazones as anthelmintic agents. PCT Int Appl 1987;WO 8706127;A1;19871022.
32. Moll J, Knapp S, Dalvit C, Trosset JY, Sundstrom M, Mantegani S. Preparation of heterocyclic compounds which interact with beta-catenin/TCF-4 binding site. PCT Int Appl 2003;WO 2003006447;A2;20030123.
33. Lepourcelet M, Chen YNP, France DS, Wang H, Crews P, Petersen F, Bruseo C, Wood AW, Shivdasani RA. Small-molecule antagonists of the oncogenic Tcf/β-catenin protein complex. Cancer Cell 2004;5:91–102.