Probing the druggability of protein–protein interactions: targeting the Notch1 receptor ankyrin domain using a fragment-based approach

Noha Abdel-Rahman*,1, Alfonso Martinez-Arias† and Tom L. Blundell*2
1Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, U.K., and †Department of Genetics, University of Cambridge, Cambridge CB2 3EH, U.K.

Abstract
In order to achieve greater selectivity in drug discovery, researchers in both academia and industry are targeting cell regulatory systems. This often involves targeting the protein–protein interactions of regulatory multiprotein assemblies. Protein–protein interfaces are widely recognized to be challenging targets as they tend to be large and relatively flat, and therefore usually do not have the concave binding sites that characterize the so-called ‘druggable genome’. One such prototypic multiprotein target is the Notch transcription complex, where an extensive network of protein–protein interactions stabilize the ternary complex comprising the ankyrin domain, CSL (CBF1/suppressor of Hairless/Lag-1) and MAML (Mastermind-like). Enhanced Notch activity is implicated in the development of T-ALL (T-cell acute lymphoblastic leukaemia) and selective inhibitors of Notch would be useful cancer medicines. In the present paper, we describe a fragment-based approach to explore the druggability of the ankyrin domain. Using biophysical methods and X-ray crystal structure analyses, we demonstrate that molecules can bind to the surface of the ankyrin domain at the interface region with CSL and MAML. We show that they probably represent starting points for designing larger compounds that can inhibit important protein–protein interactions that stabilize the Notch complex. Given the relatively featureless topography of the ankyrin domain, this unexpected development should encourage others to explore the druggability of such challenging multiprotein systems using fragment-based approaches.

Targeting protein–protein interactions
The difficulties of gaining selectivity in therapeutic intervention by targeting enzymes such as proteases or protein kinases in regulatory pathways have become more evident over the last decade. Attention is now turning to protein–protein interactions that provide an important contribution to the regulation of most biochemical pathways involved in cell signalling, growth and survival [1]. The Notch pathway provides an excellent example of these challenges [2].

Molecular complexes of the kind involved in the intracellular events in Notch signalling tend to have protein–protein interfaces that are often defined as ‘undruggable’. They are generally large (~1500–3000 Å² [1 Å = 0.1 nm]), flat and poor in distinguishing features [3] making the design of small-molecule antagonists a difficult task [4–6]. Furthermore, many contact surfaces involve residues that are disordered in the uncomplexed polypeptide and undergo a disorder–order transition on assembly of the regulatory complex. Nevertheless, encouragement that targeting such complexes might be successful has come from the discovery of ‘hotspots’, residues that contribute a significant amount to the binding interaction of the interacting proteins [7]. Hotspots are usually found in the core of the interface surrounded by residues that are less important to the stability [8]. In some complexes, hotspot residues facilitate conformational change, leading to binding sites that can accommodate small molecules [1].

Several novel approaches have been pursued for targeting protein–protein interfaces. These include the design of proteomimetics, molecules that mimic elements of the protein interface such as α-helical mimetics [9,10], and the synthesis of stapled peptides, which target α-helical molecular recognition sites and have cross-links between adjacent residues to stabilize helices and resist proteolysis. MAML (Mastermind-like)–Notch1 has been a target in the development of such campaigns [2].

A more generic approach is to use fragment-based drug discovery, which allows chemical space to be explored efficiently by identifying ‘hits’ of low affinity, albeit high ligand efficiency. These can be tethered to engineered cysteine thiols, an approach pioneered by Wells and McClendon [1]. Alternatively, uncomplexed components of the multiprotein system can be stabilized and conventional fragment-based screening developed [11–13].

In the present paper, we describe the use of untethered fragment-based approaches to targeting protein–protein
interactions. We have exploited biophysical and structural approaches to investigate fragment binding to the Notch ankyrin domain in an attempt to explore the druggability of this target with a view to discovering new leads of therapeutic value.

The Notch receptor
The Notch signalling pathway is one of several that enable cell communication during the development process [14–16]. It has proved an archetypal area for investigating functional, genetic and structural aspects of a major signalling pathway. Abnormalities associated with the signalling events are linked with a variety of diseases and cancer. The first evidence of a link of the mutations in Notch to cancer was found in T-ALL (T-cell acute lymphoblastic leukaemia) patients, of whom 10% had chromosomal translocations that resulted in expression of the Notch receptor from which the extracellular domain was absent [17].

The Notch receptor is a transmembrane protein consisting of extracellular and intracellular regions [18,19]. The extracellular region is formed of 36 EGF (epidermal growth factor) repeats and three Lin12 Notch repeats. The intracellular region includes an ankyrin domain consisting of seven ankyrin repeats. Site-directed mutagenesis experiments demonstrate that the ankyrin domain is instrumental in the signalling activity of Notch [18,20,21].

Signalling is initiated when a ligand of the DSL (Delta, Serrate and Lag2) family in one cell interacts with the extracellular domain of Notch receptor in an adjacent cell. This results in two sequential cleavage processes; the first is mediated by a member of the ADAM (a disintegrin and metalloprotease) family of metalloproteases that creates a primed substrate for the action of γ-secretase complex [22,23]. The NIC (Notch intracellular domain) is then released and translocated to the nucleus where it complexes with CSL (CBF1/suppressor of Hairless/Lag-1) transcription factor, forming a groove [24]. The MAML protein is accommodated within this groove to form an active ternary transcription complex [25], which induces transcription of target genes.

Targeting the Notch1 receptor
Different strategies have been designed to target the Notch pathway. γ-Secretase inhibitors were developed as potential therapeutic agents and reached the early stages of clinical trials [26]. However, they suffered from a lack of selectivity in addition to their gastrointestinal toxicity, which rendered them less attractive. Antibodies have also been developed either to interfere with ligand binding or to hinder the cleavage step by stabilizing an autoinhibition conformation of Notch receptor [27–29].

Targeting the downstream events of the Notch pathway may represent an alternative approach to address the problem of selectivity. Interfering with the assembly of the ternary Notch complex would require inhibiting an array of protein–protein interactions. The ankyrin domain acts a scaffold for these interactions engaging mostly in electrostatic interactions. The ankyrin surface is flat with shallow pockets, making it difficult to target by employing traditional HTS (high-throughput screening) methods. We have carried out some preliminary experiments using a structure-guided fragment-based approach that gives hope of developing ligands and, possibly, drug candidates targeted at the ankyrin domain of the Notch receptor.

Fragment screening of the ankyrin domain
We assembled 1201 fragments mainly with libraries of compounds from a commercial supplier (Maybridge) that were rule-of-three compliant [30]. The library was not target-tailored, but was reasonably diverse. It was screened against the ankyrin domain using a fluorescent-based thermal shift assay [31] and positive hits were identified as those that caused a shift in thermal unfolding temperature ($T_{m}$) by 0.5 °C or more. This was taken to indicate that the fragment was binding and stabilizing the protein. The final number of hits was 36 from a library of 1201 compounds, giving a hit rate of 2.99% (≈3%). This value is comparable with the hit rate (3.25%) with the same fragment library when screening enzyme targets in our laboratory [32]. The hit rate can give an indication of the druggability of the target [33]. Although protein–protein interfaces have been emerging as targets for drug discovery, their druggability is expected to be lower than other traditional targets when using HTS libraries for screening due to fact that large molecules are difficult to accommodate in small cavities found at protein–protein interfaces. However, smaller fragments can occupy these small pockets and provide an excellent starting point for growing fragment hits to leads (D.E. Scott, M.T. Ehebauer, T. Pukala, T.L. Blundell, A.R. Venkitaraman, C. Abell and M. Hyvönen, unpublished work). Experimental details can be found in the Supplementary Online Data at http://www.biochemsoctrans.org/bst/039/bst0391327add.htm.

The positive hits were classified according to their structural features into five groups (Figure 1). The first subgroup consisted of 13 fragments with a common phenyl ring linked through one atom to a heterocyclic five- or six- or even seven-membered ring, which could be either aliphatic or aromatic. The preference for binding to this scaffold is likely to be dependent on a degree of flexibility inferred by the one-atom bridge. The compound that contained a seven-membered 1,4-diazepan ring was found to have the highest $\Delta T_{m}$ within this group, suggesting that the 1,4-diazepan ring could be a favourable scaffold. The second subgroup consisted of six bicyclic fused rings where the phenyl ring is the common feature. The $\Delta T_{m}$ values were comparable, hence it was difficult to rank them. It appeared that there is no preference for a particular bicyclic pair of rings. The third group consisted of 14 biaryl compounds, where at least one ring is a phenyl or pyridyl ring. These compounds are similar to biphenyl systems, which were found to be the preferred moiety to bind to proteins as shown in a previous study.
Classification of fragment hits according to their distinctive chemical structures

Five main classes were found to bind by fluorescent-based thermal-shift screening. These include: (a) benzyl derivatives, (b) fused bicyclic rings, (c) biaryl compounds, (d) phenyl derivatives, and (e) five-membered heterocyclic rings.

Interactions are shown in (A) with Asp$^{2109}$, Glu$^{2076}$ and Tyr$^{2075}$ and in (B) at the dimer interface involving Asn$^{1984}$ and His$^{2019}$.

The biaryl ring systems may behave like biphenyls, which can be involved in many interactions with proteins due to their flexibility that allows them to be accommodated at protein surfaces [35]. The phenyl rings were found to be the least preferred moiety as shown previously [35]. However, the fourth group included ten phenyl derivatives and a piperidine ring. The size and shape of phenyl and six-membered rings may be more suitable for binding to shallow protein surfaces and targeting protein–protein interfaces that tend to be more lipophilic, even with fewer interactions than larger and more complex ring systems. The final subgroup is the smallest and least complex, consisting of five-membered rings. It includes only five compounds, with the imidazole ring being the most favourable. These small rings might be mimicking the side chain of histidine residues and be capable of binding to shallow and flat pockets across the interface. Although this screen could be considered a yes/no answer screening method, especially in the absence of binding enthalpy values, it can give an indication of the type of molecular scaffolds that the protein target prefers to bind.

Structural characterization of ankyrin fragment complexes

Several campaigns of soaking and co-crystallization of the positive fragment hits with the ankyrin domain resulted in no significant electron density for the majority of the fragments, probably due to low binding affinities and consequently low occupancies. However, in two ankyrin fragment crystal structures, determined at 2.4 and 2.5 Å resolutions, fragments could be fitted into convincing electron density contoured at 1σ in ways that gave clearly defined and plausible interactions with the ankyrin domain amino acid residues. The fragments were each found binding at two separate locations. The first fragment, designated 12C05, was found in one site, positioned between the upper helices of repeats six and seven (Figure 2A), apparently stabilized by weak hydrogen bonds, a salt bridge and a π–π interaction. The second site occupied by fragment...
12C05 was between the two protein chains in the asymmetric unit (Figure 2B). The interactions are mainly weak hydrogen bonds with residues Asn1984 in chain A and His2109 in chain B, and one interaction mediated through a water molecule. Similarly, the second fragment 9F07 was found binding to two distinct sites, the first located between the upper helices of repeats five and six (Figure 3A). Although the electron density was weak (1σ), it was possible to fit the fragment to the electron density. The fragment binds at a solvent-exposed shallow site, in which the nitrogen atom of the dihydroisoquinoline ring of the fragment forms a hydrogen bond with the amino side chain group of Asn2040. At the second site between two ankyrin chains in the asymmetric unit (Figure 3B), the fragment forms one hydrogen bond with His2019 between the NH of the fragment and the nitrogen of the imidazole ring of His2019 and is stacked between the two chains through hydrophobic interactions.

**Implications for complex formation**

The crystal structure of the ankyrin domain in complex with the fragments was superposed on that of the ankyrin domain in complex with MAML and CSL (PDB code 2F8X) [36]. The first site where 12C05 is binding to chain A was found to be at the interface between ankyrin and MAML. In fact, the pyrrolidine ring was found at the position where Arg40 of MAML lies at the interface with ankyrin where it interacts with Gly2073 and Asp2109 of ankyrin through its guanidinium group. Indeed the nitrogen atom of the pyrrolidine ring mimics the guanidinium group of Arg40 (Figure 4). Mutational studies have been carried out to analyse the protein–protein interactions that are involved in complex formation [37,38]. A single mutation of Asp2109 to alanine did not disrupt the complex formation. The authors reached the conclusion that the stabilization of the complex depends
Implications for dimerization

The crystal structure contains two ankyrin molecules forming a putative dimer in the asymmetric unit [39], where interactions between the protomers are more extensive than those with other molecules in the crystal lattice. There have been conflicting views on the significance of dimerization of the ankyrin domain. It was first assumed that a dimeric ankyrin domain of Drosophila was merely a feature of the crystal structure and had no physiological significance, as it was found as a monomer in solutions of different pH values and in different ionic strengths [40]. However, yeast interaction-trap assays have demonstrated that the ankyrin domain in Drosophila could be involved in homotypic interactions [41]. This confirmed earlier studies on the Glp-1 homologue of Caenorhabditis elegans that demonstrated a homotypic interaction using constructs encoding the ankyrin domain through yeast two-hybrid assays [20]. Dimerization of human Notch1 ankyrin in solution was also observed from SAXS (small-angle X-ray scattering) data [42]. Analysis of SAXS data for the Notch RAMANK (molecule consisting of the RAM region and the ankyrin domain) indicated a monomeric form [42]. This suggested that dimerization is characteristic of the ankyrin domain only when expressed on its own and that it was unlikely to be of any biological relevance in complex formation and gene transcription. However, previous studies have investigated the role of dimerization in transcription of Notch target genes [43]. These studies were based on evolutionary conserved regions in CSL-binding sites and on crystal contacts observed between ankyrin molecules. The contact residues between the two chains are engaged mainly through electrostatic interactions. Arg1985, Lys1946 and Glu1950 were found to be essential for dimerization, and single mutations of these residues led to inhibition of transcription [43]. This has been reinforced by the recent crystallographic structure of a dimeric Notch1 complex bound to HES1 paired-site DNA [44].

The fragment 12C05 was found at the dimerization interface binding to Asn1984 adjacent to Arg1985 (Figure 2B). The fragment did not interfere with any of the electrostatic interactions that stabilize the two ankyrin domain chains. However, it should be possible to grow the fragment in order to optimize its interactions with the guanidinium group of Arg1985; this could potentially inhibit the dimerization of the assembled complex. Further optimization or fragment growing could provide a starting point for a molecule that diminishes this specific interaction. Similarly, fragment 9F07 was found at the dimerization site binding to the same His2019 as fragment 12C05 through the same nitrogen of the imidazole ring (Figure 3B). This suggests that this site is a possible target to explore in order to interfere with the transcription process.

Fragment-binding affinities and implications for biological activity

Surface plasmon resonance was used to determine the binding affinities of the fragments, which were in
the millimolar range as expected. Fragment 12C05 could be fitted to a 1:1 kinetic model (Figure 6), despite its binding in two distinctive sites, implying that fragment is very weakly binding in one location. The $K_d$ value of the fragment 12C05 was 7.88 mM, whereas the fragment 9F07 bound very weakly at values >40 mM. Hence, these fragments could not be tested in biological assays at these concentrations, particularly because that these fragments are dissolved in DMSO. The affinities of different Notch polypeptides to the CSL–DNA complex were in the micromolar range, as reported previously [37]. Thus inhibiting the Notch transcription in reporter assays would require designing larger compounds with higher affinities using these fragments as starting points.

**Future developments on targeting protein–protein interactions**

The results reported in the present paper show the first structures of small molecules binding to the ankyrin domain of the Notch1 receptor. These fragments may be considered starting points for the design of larger compounds, which can be a challenging step. Further work is now needed to grow the fragments, perhaps by docking a set of commercial compounds selected according to a retrospective substructure search based on the initial fragment. Fragment growing needs to be studied extensively and different approaches and strategies explored.

The interface of the ankyrin domain involved in protein–protein interactions is a relatively flat surface devoid of distinct pockets; it has the features of an ‘undruggable’ binding surface, and most algorithms used in the pharmaceutical industry would consider it unlikely that it could be the basis for effective drug design [4]. Furthermore, the ankyrin domain does not undergo significant conformational changes upon forming the Notch transcription complex. Indeed, there is no evidence that it has hotspots that allow an induced fit of a small chemical ligand, as found in some successful exercises in targeting protein–protein interfaces [1]. Our results with fragment binding contradict these widely held assumptions and indicate that domains that are flat and do not show induced fit may still be druggable.

Further understanding of protein–ligand recognition would assist in predicting the druggability of such protein–protein interfaces. Targeting the Notch transcription complex has proven successful by designing synthetic peptides mimicking the MAML. However, we have described a more generic approach to target protein–protein interactions, employing a fragment-based approach rather than peptidomimetics.

**Acknowledgements**

We thank Matthias Ehebauer for his helpful discussion.

**Funding**

N.A. was sponsored by BP-Egypt and Cambridge Overseas Trust. T.L.B. thanks the Wellcome Trust for support through a Programme Grant [grant number 079281/Z/06/Z].

**References**


34 Reference deleted


Received 27 May 2011
doi:10.1042/BST0391327
SUPPLEMENTARY ONLINE DATA

Probing the druggability of protein–protein interactions: targeting the Notch1 receptor ankyrin domain using a fragment-based approach

Noha Abdel-Rahman*, Alfonso Martinez-Arias† and Tom L. Blundell*2
*Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, U.K., and †Department of Genetics, University of Cambridge, Cambridge CB2 3EH, U.K.

Protein expression and purification
The ankyrin domain of Notch1 receptor was produced as reported previously [1]. The protein was purified in a buffer constituting of 50 mM NaCl and 50 mM Tris/HCl (pH 8) for further experiments.

Fragment screening by fluorescent-based thermal shift assay
A fragment library was assembled by purchasing fragments from a commercial supplier (Maybridge) and including another 100 compounds prepared in the Chemistry Department, University of Cambridge, in the laboratory of Professor Chris Abell. The fragments were dispensed in 96-well thin-walled plates (Bio-Rad Laboratories) from a 100 mM stock in neat DMSO to form 5 mM solution in a 100 μl volume. The protein was mixed at 10 μM with SYPRO® Orange dye (1:250) in a buffer consisting of 10 mM Chex [2-(N-cyclohexylamino)ethanesulfonic acid], 400 mM NaCl and 5 mM DTT (dithiothreitol). Controls were run in the same plate using DMSO instead of the fragment. The plates were then heated in an iCycler iQ Real Time PCR Detection System (Bio-Rad Laboratories) from 25 °C to 60 °C in 0.5 °C increments. Fluorescence was detected using a CCD (charge-coupled device) camera. SYPRO® Orange dye has fluorescence excitation and emission wavelengths of 490 and 525 nm respectively.

Crystallization and structural determination
The protein was crystallized at a concentration of 10 μM using the hanging-drop method. The crystallizing solution was composed of 0.7 M (NH4)2HPO4, 0.2 M NaCl and imidazole (pH 8). Crystals were then soaked in 100 mM fragment solutions in the mother liquor for 4 h. Datasets were collected at ESRF (European Synchrotron Radiation Facility), Grenoble, France, and processed using MOSFLM then scaled using SCALA program, a part of the CCP4 package. The structures were solved by molecular replacement using the program PHASER and the co-ordinates of the Notch ankyrin domain (PDB code 1YYH) as a probe. Several rounds of refinement were performed using REFMAC5. The σA-weighted 2Fo −Fc and Fo −Fc electron-density maps were visualized to allow rebuilding and refitting using Coot. Refinement statistics are shown in Supplementary Table S1.

Measuring binding affinities using surface plasmon resonance
The ankyrin domain was immobilized on a CM5 chip using HBS (Hepes-buffered saline: 10 mM Hepes, pH 7.4, 150 mM NaCl and 0.05% surfactant P20) as a running buffer. Serial dilutions of the fragments were prepared in HBS. Depending on the solubility, some fragments were dissolved in 5% DMSO which was added in the running buffer. Solvent correction was applied in this case. The Biacore T100 system was used and all runs were carried out at 25 °C in triplicate.

1Present address: Department of Biochemistry, Faculty of Pharmacy, Mansoura University, Mansoura, Egypt.
2To whom correspondence should be addressed (email tlb20@cam.ac.uk).

Table S1 | Data collection and refinement statistics ankyrin fragment structures

Values in parentheses are for the highest resolution shell. $R_{\text{sym}} = \Sigma_{i} |I_{i} - \langle I \rangle| / \Sigma_{i} I_{i}$, where $I_{i}$ is the intensity of reflection $h$ and $\langle I \rangle$ is the mean intensity of all symmetry-related reflections. $R_{\text{cryst}} = \Sigma_{i} |F_{\text{obs}}| - |F_{\text{calc}}| / \Sigma_{i} F_{\text{obs}}$, where $F_{\text{obs}}$ and $F_{\text{calc}}$ are observed and calculated structure factor amplitudes respectively. $R_{\text{free}}$, as for $R_{\text{cryst}}$ using a random subset of data excluded from the refinement, 5 % of the total dataset was used. Estimated co-ordinate error based on $R$ value was calculated using Refmac. RMSD, root mean square deviation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fragment 12C05</th>
<th>Fragment 9F07</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>$P6_5$</td>
<td>$P6_5$</td>
</tr>
<tr>
<td>Unit cell</td>
<td>$a=b=97.94$ Å; $c=109.44$ Å; $\alpha=\beta=90^\circ$; $\gamma=120^\circ$</td>
<td>$a=b=97.43$ Å; $c=110.23$ Å; $\alpha=\beta=90^\circ$; $\gamma=120^\circ$</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>21.20-2.53 Å (2.40-2.40 Å)</td>
<td>33.69-2.42 Å (2.30-2.30 Å)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.18</td>
<td>99.98</td>
</tr>
<tr>
<td>$R_{\text{sym}}$ (%)</td>
<td>5 (27.5)</td>
<td>3.5 (22.20)</td>
</tr>
<tr>
<td>Average $I/\langle I \rangle$</td>
<td>6.6</td>
<td>7.3</td>
</tr>
<tr>
<td>Number of unique reflections</td>
<td>21 935</td>
<td>22 060</td>
</tr>
<tr>
<td>Wilson plot $B$ factor</td>
<td>34.53</td>
<td>50.93</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_{\text{cryst}}$</td>
<td>22.39</td>
<td>24.3</td>
</tr>
<tr>
<td>$R_{\text{free}}$</td>
<td>29.94</td>
<td>31.3</td>
</tr>
<tr>
<td>Molecules per asymmetric unit</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Number of ligand molecules</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Estimated co-ordinate error (Å)</td>
<td>0.30</td>
<td>0.31</td>
</tr>
<tr>
<td>RMSD bonds (Å)</td>
<td>0.024</td>
<td>0.026</td>
</tr>
<tr>
<td>RMSD angles (°)</td>
<td>2.17</td>
<td>2.42</td>
</tr>
</tbody>
</table>

Reference


Received 27 May 2011
doi:10.1042/BST0391327