Stefano Mangani *Editor*

Disruption of Protein–Protein Interfaces

In Search of New Inhibitors

Springer

Disruption of Protein–Protein Interfaces

Stefano Mangani
Editor

Disruption of Protein–Protein Interfaces

In Search of New Inhibitors

Springer

Editor
Stefano Mangani
Department of Biotechnology, Chemistry and Pharmacy
University of Siena
Siena
Italy

ISBN 978-3-642-37998-7 ISBN 978-3-642-37999-4 (eBook)
DOI 10.1007/978-3-642-37999-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013942138

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The large size and variety of the human protein interactome and the obvious relevance of the protein–protein interactions in every physiological function, render protein–protein interactions at the same time an extremely challenging and attractive target for developing of new therapeutic substances. A further interesting aspect of targeting protein–protein interactions (PPI) for drug discovery is that, at least in some cases, molecules directed against PPI may provide a way to overcome the resistance mechanisms encountered for active site binding enzyme inhibitors. The considerations above can be extended to bacterial and viral interactome, further expanding the vastness and complexity of the subject of PPI inhibition and modulation.

For these reasons, there is no surprise that targeting PPI has become a subject of intense research activity in both industry and academia over the past decade. Reviewing the literature and the available databases shows that over 150 small molecule compounds have been found to inhibit about 20 PPI targets and that some of these molecules have already reached, or are about to reach, the drug market. As usual, these figures can have a double reading: the optimistic interpretation is that these successes demonstrate the validity of the approach, while the critics might say that the huge effort dedicated to find PPI inhibitors has resulted in scarce results. The fact is that exploring PPI is a stimulating new subject of study that is relevant for the basic knowledge on the chemistry of living organisms with the important outcome to offer the possibility of opening a new era of drug discovery.

This book is a collection of essays from Italian research groups from Industry and University involved in drug discovery and, although the book presents different subjects in each chapter, the unifying idea comes from our belief that only an integrated approach of the different techniques nowadays available, may overcome the challenges presented by this new frontier in drug discovery. For this reason,

the book opens with reviews about the current status of the research on PPI in drug discovery and goes on by presenting the state of the art in basic techniques like computational tools, NMR, X-ray crystallography and FRET that, integrated, may give the opportunity of success in this field.

Siena, March 2013 Stefano Mangani

Contents

Chapter 1
Drug Discovery by Targeting Protein–Protein Interactions

Laura Bettinetti, Matteo Magnani and Alessandro Padova

1.1 Introduction

For many years, drug discovery's main interest has focused on protein–ligand interactions, such as enzyme inhibition or transmembrane receptor modulation. Drug discovery approaches were based on endogenous ligand knowledge as starting point for small molecule target modulation. This approach allowed the identification of small molecules with good ADMET profile thanks to the elaboration of best practices together with the development of high-throughput technologies.

Over the last decade, a new class of molecular targets captured the attention and became the new frontier of drug discovery: the protein–protein interactions (PPIs). Protein interactions regulate many cellular functions, such as DNA replication, transcription, translation, and transmembrane signal transduction. At the same time, the alteration of PPI physiological functions is implicated in several diseases. For this reason, the interest on PPIs as new molecular target in drug discovery has become a reality resulting in a surge of new potential therapeutic targets.

L. Bettinetti · M. Magnani · A. Padova (✉)
Drug Design Unit, Department of Molecular Informatics, Siena Biotech S.p.A, Strada del Petriccio e Belriguardo 35, 53100 Siena, Italy
e-mail: a.padova@irbm.it

Present Address:
A. Padova
Strategy & Alliances, IRBM Promidis S.r.l., Via Pontina Km 30,600 00040 Pomezia (RM), Italy

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4_1,

1.2 Basic Principles of Protein–Protein Interactions

PPIs can give origin to obligatory (or permanent) complexes, when these are constituted by proteins whose function is exerted only when bound to the partner protein. Permanent PPIs are usually tight, with strong hydrophobic effect, well packed and with few water molecules trapped inside. However, PPIs can give origin also to transient complexes. The formation of the latter class depends on the functional state of the involved proteins, the interfaces are mainly less extensive and more polar/charged, and the association is weaker than in permanent complexes [1]. Obligatory complexes are usually homodimeric structures (dimers formed by identical subunits), but there are also examples of heterodimeric obligatory complexes. Several examples have been classified by Subhajyoti and coworkers [2]. For example, the interaction of cytochrome c/cytochrome c peroxidase represents a transient complex, while interleukin-1 beta convertase interacts with itself giving an obligatory complex.

Some proteins involved in PPI do not show a stable and well-defined structure under physiological conditions, but give origin to a well-ordered complex only after interaction with their partner. In this case, we can talk of disordered or intrinsically unstructured proteins, which have a broad range of functions, for example, transcription, translation, and cellular signaling. Interestingly, the disordered state of these proteins closely resembles their denatured state, and thus, the knowledge of the denatured state makes possible to understand the folding process [1].

Proteins that interact with only one or few other proteins are termed loner proteins, while those interacting with several partners are known as hub proteins

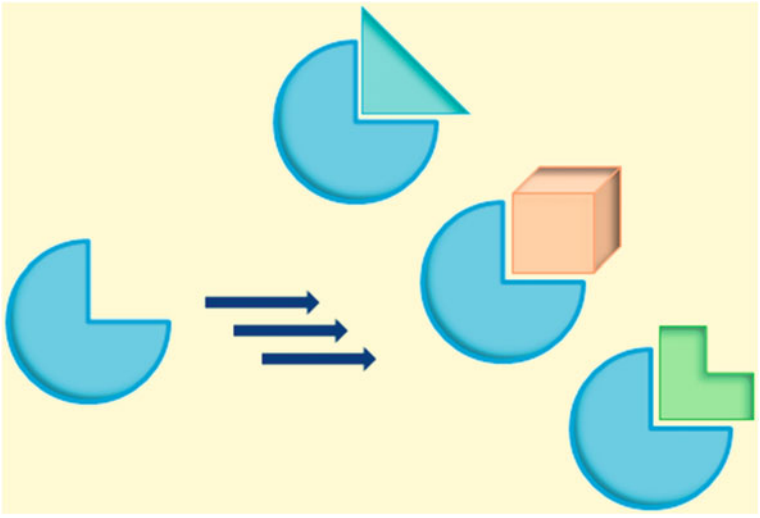


Fig. 1.1 Hub proteins have highly versatile interfaces which allow interaction with several partners, while maintaining the same role in all complexes

[3]. Hub proteins have interfaces of very different size, with structural and/or sequence repeats, and are at least partially disordered (Fig. 1.1). They have a high versatility but mainly explain the same function for different partners.

According to a common assumption in drug discovery, PPIs contact surfaces have been generally depicted as a large (1,500–3,000 $Å^2$) and almost flat landscape with no or few crevices suitable as small molecule binding sites [4]. However, Keskin and coworkers reviewed this 'flat' approach, discussing about pockets, crevices, and indentations which are occupied by water in the unbound state [1]. The interaction with the partner protein is able to replace completely or partially the water in the involved pockets, called complemented pockets, while it has no effect on the unfilled pocket. It is now established that in many protein–protein interfaces, a restricted portion (around 600 $Å^2$) of the entire contact surface is responsible for most of the binding energy. This portion is referred to as "hot spot" and represents the most attractive region to interact with for small molecules aimed to efficaciously hamper the interaction between proteins [5, 6]. The hot spot generally consists of a rim region, similar to the rest of the protein, and a core region, usually rich of aromatic residues responsible for key hydrophobic interactions. Even if it has been observed that often protein–protein interfaces have preferred architecture [1], the exact definition of available hot spots can be derived only from labeled protein NMR and crystallographic experiments [7]. These studies are limited by protein size and of course by ease of protein expression and purification. This is the reason why in the last few years increasing efforts have been made to develop computational approaches capable to identify and analyze protein surface hot spots [8–10].

Interactions between proteins are mainly guided by hydrophobic contacts, though hydrogen bonds, electrostatic interactions, and covalent bonds can also be involved. When dealing with PPIs, it is important to be aware that proteins interact in a dynamic way, via a dynamic formation and dissociation of the complex. Many proteins assume a well-defined conformation only after binding, while they exist in a disordered state in solution. In these cases, standard techniques such as virtual screening and structure-based design, when relying on X-ray or NMR crystal structures of isolated proteins, might not be viable during the hit identification phase [1].

Small molecules can modulate PPIs by interacting not only with the protein–protein interface region, but also with allosteric sites, located in regions of proteins not directly involved in the interaction with the partner (Fig. 1.2) [11]. The rearrangement originated by the ligand binding is an alteration of the normal protein conformation, affecting also the binding region which thus becomes inaccessible to the partner protein. Allosteric modulation can play a crucial role in both hub and loner proteins. This avenue has different advantages, including high specificity and the possibility to have multiple allosteric sites for the same protein.

The classical drug discovery approaches such as high-throughput screening and fragment screening have been relatively useful in finding PPIs inhibitors. Several methods to identify PPIs have been developed during the last two decades, including high-throughput, fragment-based, in vitro and in silico screening [4]. Moreover, the large interest in the field is leading to the generation of improved

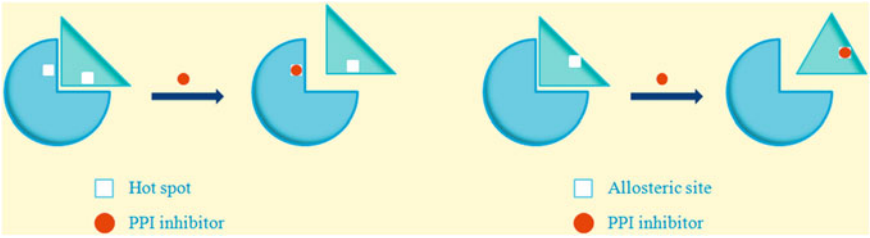


Fig. 1.2 *Hot spots* and *allosteric sites* in protein–protein interactions

technologies allowing for the high-throughput study for protein–protein interactions in several model organisms [12]. However, despite some successful examples demonstrating the feasibility of PPIs modulation for therapeutic intervention, the search for PPIs inhibitors is still challenging for several reasons, most of them being ascribable to the nature of protein–protein interfaces and available screening technologies. Indeed, although the extent and flatness of surfaces can be partially overcome by exploiting the presence of hot spot regions, the hydrophobicity that usually characterizes the binding sites requires compounds with particularly high shape complementarity to be active.

Even more important, interaction with hydrophobic sites usually implies the abundant presence in ligands of lipophilic moieties. Although these moieties can guarantee high binding energies, they negatively affect important physico-chemical properties of compounds (in particular solubility), thus generating serious issues about their developability.

Another aspect that characterizes protein–protein interfaces is that, unlike canonical drug targets, they did evolve to bind proteins and not small molecules. This makes possible the existence in binding regions of steric and geometric requirements difficult to be fully satisfied by the limited degrees of freedom of small molecules.

The dynamic behavior of protein surfaces can also have a role in hindering the identification of PPIs inhibitors. When protein-binding sites are characterized by high flexibility, indeed, X-ray or NMR structures commonly used for rational drug design can give only a partial picture of real scenario, since there may be several protein conformations still undisclosed by experimental methods. In these cases, computational methodologies can be a useful support to simulate protein flexibility and provide pools of suitable conformations for structure-based design of ligands [13].

A further challenge in this field is how to select druggable and disease modifying PPIs among the thousand available PPIs. Several in vitro and in vivo methods have been developed to identify possible PPIs, but how can protein druggability be predicted? Computational approaches have been useful to find the ligand-binding site analyzing the protein surface but have not been able yet to define whether a small molecule will be capable of interacting with high affinity and specificity. Biochemical screening with a very broad and diverse compound collection represents another useful way to establish druggability of PPIs and find positive hits. When a

target yields several and diverse hits, this is assumed to be highly druggable, whereas near zero hit rate, the target is deemed not druggable. However, there are several drawbacks with this approach: first of all, it requires a very lengthy and expensive process and it is generally limited to bimolecular PPIs rather than multiprotein complexes. Secondly, it is fundamental to have access to a very broad compound collection (hundreds of thousands compounds) in order to have an exhaustive chemical diversity that will hopefully deliver a molecule that will match the target biological space; moreover, high- and medium-throughput screening data are generally plagued by false positives due to several reasons such as compound contamination with active impurities or possible cross-activation of the assay readouts, thus requiring usage of an orthogonal assay to eliminate false positives.

One of the most successful approaches applied so far to identify PPIs is NMR-based screening of fragment libraries. Indeed, 2D heteronuclear correlation spectra have provided reliable results limiting the false positives/negatives issue, while fragment-based screening has allowed to significantly reduce the number of compounds necessary for a screening campaign.

Charge, hydrophobicity, and shape have been outlined as fundamental parameters to define the druggability indices of a target combining analyses of NMR data and the knowledge of the ligand-binding site. These indices allow classification of druggability in "high," "medium," or "low" [7, 9, 14, 15].

The ligand efficiency coefficient, defined as the ratio between ΔG and the number of non-hydrogen atoms of the compounds (LE = (ΔG)/N where ΔG = − RTlnKi) [16], is widely applied in drug discovery to describe ligand–protein interactions. Interestingly, when evaluated in the field of PPIs, ligand efficiency values were found to be lower than for common kinase inhibitors, but similar to many protease inhibitors [17]. In more detail, ΔG for PPIs was estimated to be around 0.24 kcal/mol, thus implying that a compound with $K_d = 10$ nM should have a MW near 645 Da, higher than typical orally available drugs.

1.3 Strategies to Interfere with PPIs

Three main classes of protein–protein interaction ligands can be delineated so far: (1) small molecules, (2) peptides and peptidomimetics, and (3) humanized monoclonal antibodies (MAbs).

1.3.1 Small Molecules

Small molecule PPI modulators represent the most rationale approach from a drug discovery and development perspective. Many examples belonging to this class are natural products or derivatives and are characterized by physico-chemical properties violating the Lipinski's "Rule of Five," having either high molecular weight

or high logP [5, 18]. However, medicinal chemistry efforts have been able to address some of these issues delivering orally available candidates such as ABT-263 or Navitoclax (see hereinafter). In this context, it is questionable whether the Lipinski's "Rule of Five" should be applied to natural products and in general to PPI's modulators.

1.3.2 Peptides and Peptidomimetics

The second class of compounds comprise peptide and peptidomimetics, which have traditionally suffered of poor ADMET properties such as stability issues, poor tissue penetration, poor solubility, protease resistance, low oral bioavailability, and fast elimination. For these reasons, clinical peptide development has been slow for a long time in favor of small molecule approaches. Only recently, it has been shown that appropriate modifications on the simple peptidic sequence may address some of the above issues and there has been a resurgence of peptide-based approaches [19, 20].

Starting from the amino acidic sequence of the contact surface between two proteins, it may be feasible to synthesize a short peptide sequence which can hamper the interaction in a highly specific way and with low toxicity effects. Furthermore, small peptides rarely provoke immune responses with respect to monoclonal antibodies and have fewer toxicology issues than small molecules.

Several approaches are currently undergoing to develop appropriate peptides. One possibility is to start from a known or predicted protein structure and derive the sequence to interfere with its target, for example, in the design of β-sheet breaker peptides to abolish the formation of amyloid plaques in Alzheimer's disease. Combinatorial techniques are used to generate and screen peptide libraries identifying hit ligands as starting points for development. This method permits to combine the information coming from the structural approach with the use a diverse library of peptides. Finally, several other methodologies have been developed, both in cells (for example protein-fragment complementation assays, PCAs) and in vitro (Phage display, Ribosome display, mRNA display, CIS display) [12, 19, 21].

Peptide production can be achieved using recombinant techniques using bacteria, yeast or transgenic animals. Moreover, if it is necessary to modify the existing sequences, peptides may be produced in cell-free transcription/translation-based systems incorporating non-native amino acids [19].

Chemical synthesis allows for several modifications, all aimed at improving the activity and the pharmacological profile of the original peptide. The basic idea is to design structures able to mimic the endogenous secondary structure of the reference peptide and lock the biologically relevant conformation, improving stability and thus increasing half-life and oral bioavailability. Common examples of peptide modifications are introduction of amide bond surrogates (providing derivatives such as peptoids—amino acid side chain attached to the amide group—or

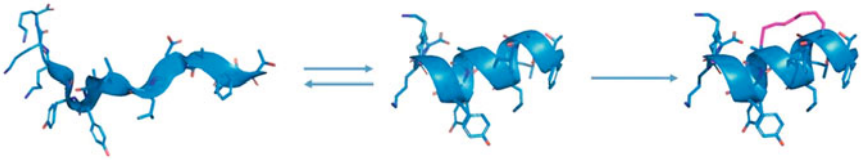


Fig. 1.3 Peptides can be "stapled" into an α-helical, biologically active conformation through hydrocarbon cross-links between non-natural aminoacids

depsipeptides—amide bond replaced by ester bonds), use of unnatural aminoacids, N-methylation of the N–H peptide group, introduction of phosphate groups, N-terminal acetylation, or C-terminal amidation [19].

Another avenue explored to create biologically stable and active peptides has been to cyclase between the N and C termini giving cyclic peptides or to covalently fix parts of the molecule to create stapled peptides. Examples of former structures occur naturally, giving peptides resistant to protease action highly potent as well as selective [22]. It is important to underline that even if the molecular masses of these cyclic peptides and other macrocycles are significantly out of small molecule developability criteria, they can possess drug-like physico-chemical and pharmacokinetic properties such as good solubility, lipophilicity, metabolic stability and bioavailability. Stapled peptides (Fig. 1.3) are particularly promising and open the chance to inhibit targets considered undruggable to date. A peptide can be "stapled" into biologically active conformation using three different approaches: helix stabilization, helical foldamers, and helical surface mimetics [23]. Helix stabilizations present side chain cross-links and hydrogen-bond surrogates which preorganize amino acid residues and initiate helix formation; helical foldamers present amino acid analogs and adopt conformations similar to those found in natural proteins, for example, β-peptides and peptoids; helical surface mimetics utilize conformationally restricted scaffolds with attached functional groups giving an α-helix.

Another avenue that has been explored in drug discovery to increase permeability and/or improve metabolic stability of peptides has been the use of D- or inverso-peptides (peptides made of D-amino acids), retro-peptides (L-peptides with reverse sequence), and retro-inverso-peptides (peptides made with D-amino acids and with reverse sequence) [19]. Examples are provided in Fig. 1.4.

1.3.3 Humanized Monoclonal Antibodies

A third class of PPI modulators includes the humanized monoclonal antibodies, which are highly target specific but with low permeability both in cell and in the brain. These are antibodies from non-human species whose protein sequences have been modified to increase their similarity to the human's one naturally produced

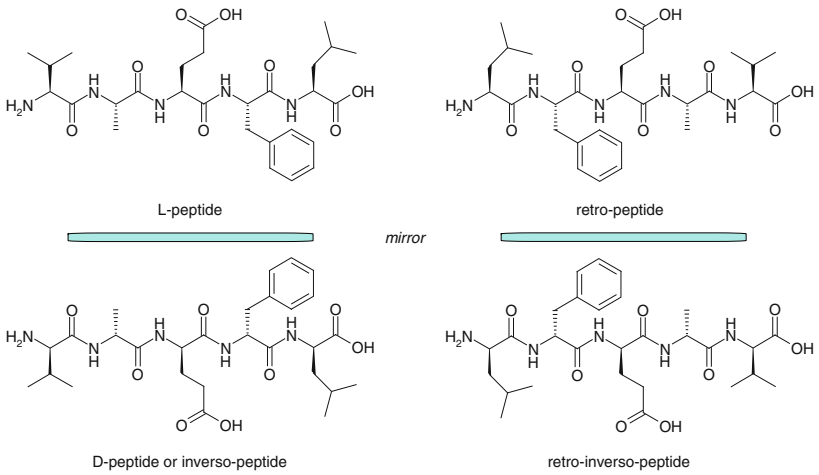


Fig. 1.4 Peptide derivatives

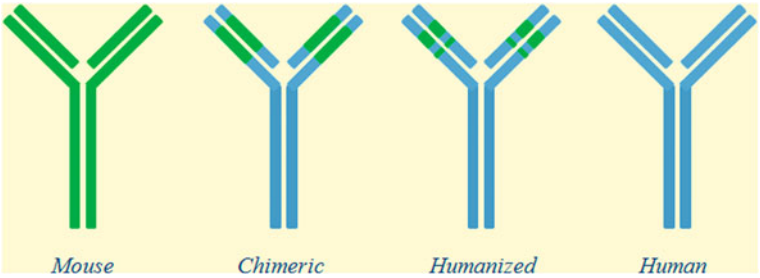


Fig. 1.5 Humanized monoclonal antibodies

(Fig. 1.5). This variation overcomes the problems of immunogenicity and inefficient secondary immune function that frequently are present in clinical use.

Monoclonal antibodies that block the interaction between interleukin 2 (IL-2) and its receptor (IL-2R) represent a successful example of targeting protein–protein interactions in immunotherapy. Protein engineering has been successful in the generation of functional antibody fragments, although alternative smaller and more compact protein frameworks are still desirable [24].

1.4 Targeting PPIs

CNS and cancer represent the two main therapeutic fields of interest in terms of PPI discovery and development [25, 26]. Several neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, and Huntington's disease are characterized by the formation of aggregate protein plaques. Inhibiting or reversing such

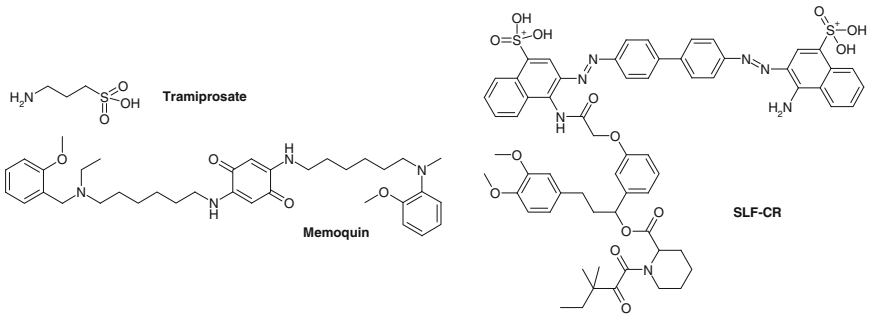


Fig. 1.6 Inhibitors of amyloid-β aggregation

aggregation process is considered a viable approach to block or revert disease phenotypes.

It has been reported that two different kinds of plaques have been observed in the brain of the Alzheimer's disease patients: amyloid-β protein aggregates and the neurofibrillary tangles of aggregated τ-protein. Examples of inhibitors of amyloid-β aggregation are tramiprosate (Alzhemed), memoquin, and SLF-CR (Fig. 1.6).

Tramiprosate (Alzhemed, Vivimind) is a PPI inhibitor that binds to monomeric amyloid-β protein disrupting its aggregation and neurotoxicity while promoting clearance from brain. Phase II trials demonstrated that the drug reduces Aβ42 in the cerebrospinal fluid of patients with Alzheimer's disease. Unfortunately, the compound failed phase III clinical trials because the study missed to demonstrate a beneficial effect on the primary outcomes, change in cognition and clinical stage [27, 28]. Memoquin is another inhibitor of the amyloid-β aggregation obtained by matching the benzoquinone fragment of the coenzyme Q10, which has antioxidant activity, and is able to inhibit the amyloid-β aggregation, with a series of cholinesterase inhibitors. This compound is reported to be orally bioavailable and to cross the blood–brain barrier [29]. SLF-CR is a bifunctional molecule designed by covalently matching a synthetic ligand for FK506-binding protein family of chaperones (SLF) and Congo red (CR). The first part of the molecules inhibits the aggregation of the amyloid-β, and the second one is able to bind the amyloid-β [30].

Huntington's disease is a neurodegenerative poly-glutamine disorder characterized by the aberrant expansion of polyQ regions in proteins. These polyQ regions cause proteins to associate and form insoluble aggregates in the brain that are common features of polyQ diseases. The exact mechanism of toxicity is not well understood; however, peptide inhibitors of aggregation could be a good therapeutic approach in the prevention and cure of Huntington's disease [31].

In Parkinson's disease, it has been shown that the overexpression of the α-synuclein protein causes the development of disease symptoms. A series of peptidomimetics and small molecules inhibiting α-synuclein aggregation were discovered. Some examples are shown in Fig. 1.7.

A common approach in targeting CNS diseases is to modulate the altered signal transduction pathway acting on the receptor involved, which is the upstream target

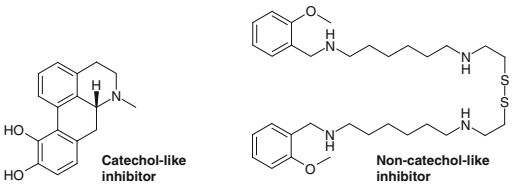


Fig. 1.7 Inhibitors of α-synuclein aggregation

of the system. In this context, G-protein-coupled receptors (GPCRs) constitute a large and diverse class of receptors widely targeted in drug discovery. GPCRs are seven-transmembrane domain receptors that interact with ligands outside the cell and activate inside signal transduction pathways through heterotrimeric G-proteins (Fig. 1.8).

When activated by a ligand, the G-protein-coupled receptor acts as a guanine nucleotide exchange factor (GEF), resulting in the dissociation of Gα and G$\beta\gamma$ subunits of G-protein. Both subunits can induce several signaling cascades. The interactions between the Gα/G$\beta\gamma$ subunits and different protein effectors are PPIs,

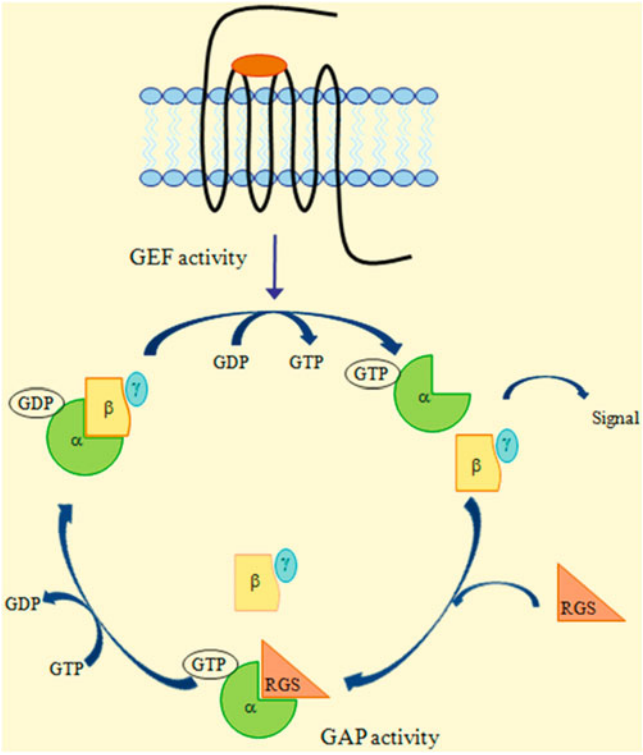


Fig. 1.8 Activation cycle of G-proteins by G-protein-coupled receptors

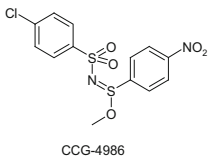


Fig. 1.9 Inhibitor of RGS4–Gα interaction

and development of selective modulators of these PPIs can modulate the signal pathway. Another approach is to modulate the activity of proteins that act as regulators of the G-protein signals (RGS). Such proteins accelerate the rate of GTP hydrolysis by Gα subunits, thus acting as GTPase-accelerating proteins (GAP activity). A number of peptides and small molecules inhibitors of RGS functions have been published [32]. One example is CCG-4986 (Fig. 1.9), the first small molecule discovered that inhibits the RGS4–Gα interaction.

In tumors, a significant number of functionally heterogeneous genes show mutations which lead to uncontrolled cell proliferation. These genes often encode hub proteins, having a strong influence on protein–protein interaction network. As a consequence, the oncology field comprises several PPIs which can be modulated for therapeutic intervention.

1.4.1 p53–MDM2

Tumor suppressor protein p-53 is known as the "guardian of the genome" for its responsibility to determine the cell fate when the integrity of its genome is damaged. Its function is to induce cell cycle arrest and apoptosis in response to DNA damage. Several human tumors present an altered p53 activity. This is due in most cases to direct mutation or deletion of the gene, which hampers its natural function, but there are also several tumors which harbor wild-type p53. In many of these tumors, p53 is downregulated by the overexpression of two oncoproteins: Murine Double Minute 2 (MDM2, HDM2 in human) and Murine Double Minute X (MDMX). These proteins bind the N-terminal transactivation domain of p53, thus holding up its transcriptional function [33]. Inducing p-53 pathway acting directly on DNA damage of its gene can also induce p53-independent pathways, causing severe toxicities in normal tissues. Conversely, recovering the impaired function of the p53 protein by disrupting the MDM2–p53 or MDMX–p53 interaction offers a better treatment for a broad spectrum of cancers.

X-ray crystallography disclosed the details of the p53–MDM2/X interaction, characterized by an MDM2/X hot spot occupied by a short α-helix of p53 (Fig. 1.10) [34, 35]. In particular, three hydrophobic residues of the p53 α-helix, namely Phe19, Trp23, and Ile26, were shown to interact with three distinct subpockets of the binding site and to play a key role for the interaction between the two proteins.

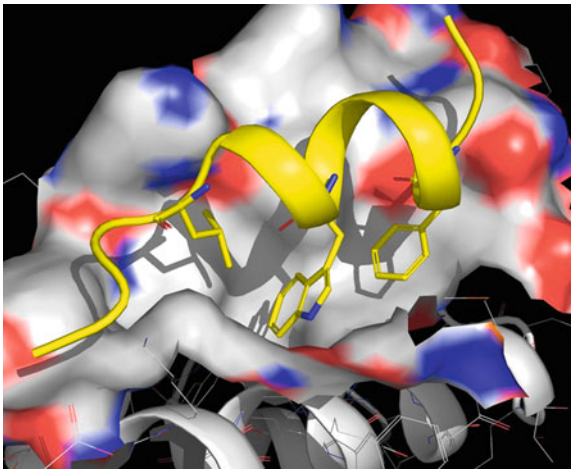
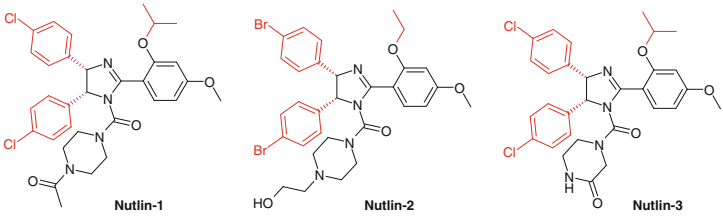


Fig. 1.10 X-ray structure of p53-MDM2 complex. p53 residues Phe19, Trp23, and Ile26, involved in fundamental hydrophobic interactions with the MDM2 hot spot, are shown as *yellow sticks*

The search for p53–MDM2 inhibitors has led to several classes of small molecules able to bind MDM2 at the p53-binding site and restore p53 activity. Reflecting the hydrophobic nature of the p53–MDM2 interaction, all of these compounds are characterized by a central scaffold that directs hydrophobic substituents toward the three MDM2 sub-pockets, thus mimicking the key side chains of p53.

One of the most important families of compounds is Nutlins from Hoffmann-La Roche, reported in the scheme below [36]:



Nutlins family. Hydrophobic moieties are highlighted in red.

The most studied compound of the family is Nutlin-3. It has a $K_i = 36$ nM toward the p53-binding site of MDM2, and in vivo experiment with Nutlin-3 and other inhibitors confirmed that small molecule inhibitors of the p53–MDM2 interaction are able to induce either cell cycle arrest or apoptosis in tumor cells, while not affecting healthy cells. The scaffold of Nutlin-3 has a tetrasubstituted imidazoline core, and it was discovered by high-throughput screening. Nutlin-1 and Nutlin-2 were, respectively, threefold and twofold less active against MDM2 than Nutlin-3. Moreover, Nutlin-2 was the first small molecule to be reported in the co-crystallized form with MDM2 [37]. The structure of the complex revealed the stringent shape complementarity between the hydrophobic substituents of Nutlin-2 and the MDM2-binding site (Fig. 1.11).

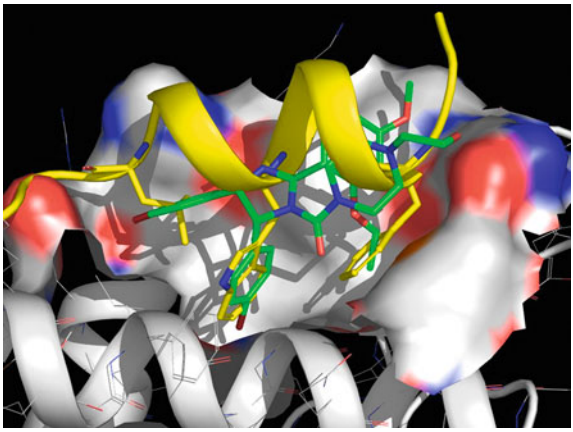


Fig. 1.11 X-ray structure of p53 (*yellow*) and Nutlin-2 (*green*) in complex with MDM2. The picture shows the good overlap between Nutlin-2 hydrophobic moieties and p53 side chains

The role of MDMX protein as a critical p53 regulator and antitumor target has emerged recently, and peptide/peptidomimetics and knockout studies confirm that an inhibitor of p53–MDMX interaction should have high therapeutic value. Nutlin-3 binds also MDMX, with K_i values in the micromolar range. Nutlin-3 has been reported to be undergoing clinical evaluation for various cancer types such as sarcoma [38], retinoblastoma [39], and lymphoma [40].

The idea that the 6-chloroxindole unit mimicked the native Trp23 configuration gave origin to MI-219 and its homologue MI-63, two representatives of a wide and well-characterized family of inhibitors. The affinities of MI-63 toward MDM2 and MDMX are 5 nM and 55 uM, respectively [41].

Another group of inhibitors reported in the literature are the benzodiazepinediones [42]. The scaffold was found through an HTS campaign of over 300,000 compounds tested. The hit MDM2-binding compound was optimized. The early lead TDP222669 has a *Ki* value of 80 nM, but suffered from low bioavailability and high clearance. Further optimization gave better physico-chemical properties with a slight loss in the binding affinity.

Another recently developed p53–MDM2 inhibitors family is the isoindolinone family. Its precursor, NU8165, had an IC_{50} around 16 μM (Fig. 1.12) [43]. Although no suitable co-crystals of isoindolinones bound to MDM2 have been obtained so far, NMR studies have provided valuable structural information. Indeed, NMR experiments using NU8165 showed a significant chemical shift change in the Leu54 region attributed to an interaction with the 3-alkoxy substituent [44]. Reasonably, additional potency may be gained by introducing rigidity to it. The binding mode model also suggested that the substitution of the N-benzyl moiety could introduce additional favorable interactions and improve potency. In fact, optimization of the precursor NU8165 gave a more potent compound, 3-(4-chlorophenyl)-3-((1-(hydroxymethyl)cyclopropyl)methoxy)-2-(4-nitrobenzyl)-isoindolin-1-one, with an IC_{50} of 0.23 μM. The resolution of the enantiomers showed that the R-enantiomer was the more potent, with an IC_{50} of 0.17 μM.

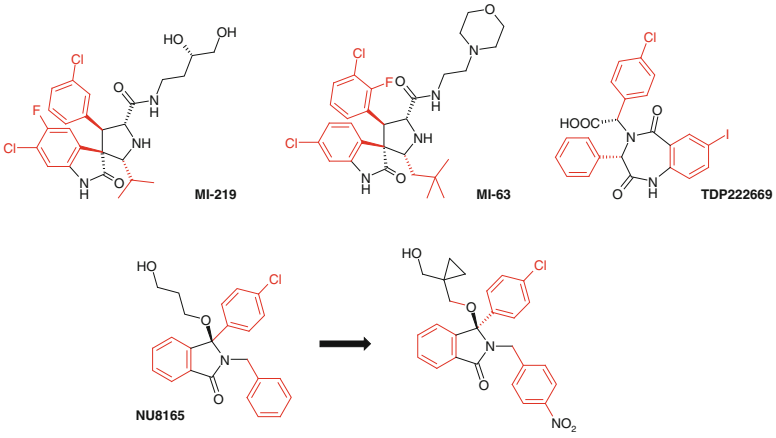


Fig. 1.12 Other families of p53–MDM2 inhibitors

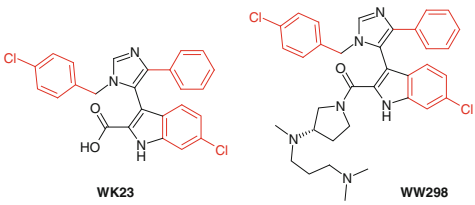


Fig. 1.13 p53–MDM2/MDMX inhibitors

A recently reported series of p53–MDM2/MDMX inhibitors are the imidazo-indoles [45]. Two compounds, namely WK23 and WW298, showed sub-micromolar *Ki* values against MDM2 (Fig. 1.13). They were also found to inhibit the p53–MDMX interaction, although with lower potency.

A large number of p53-like peptides and mini-proteins with increased binding affinity toward MDM2/MDMX compared to the wild-type p53 analog have been developed [46, 47]. All of these peptides contain the triad Phe19-Trp23-Leu26 oriented toward the binding cleft of MDM2/MDMX. It is important to notice the effect of Pro27 of the native p53 sequence, which breaks the α-helical secondary structure of the native peptide: substitution with other amino acids can increase α-helicity, finally resulting in improved peptide affinity toward the target.

Walensky and coworkers generated stabilized α-helix of p53 (SAH-p53) peptides that exhibit high affinity for HDM2, using a peptide stapling strategy in which an all-hydrocarbon cross-link was generated within natural peptides by ruthenium-catalyzed olefin metathesis. SAH-p53 treatment reactivated the p53 tumor suppressor cascade by inducing the transcription of p53-responsive genes, providing the first example of a stapled peptide that kills cancer cells by targeting a transcriptional pathway [47].

Still in the field of peptide derivatives, a retro-inverso peptide of the p53 helical peptide interacting with MDM2 was found to bind MDM2 using the same

hydrophobic residues (namely phenylalanine, tryptophan and leucine) and was proposed to adopt a right-handed helical conformation upon MDM2 binding [48]. However, it was subsequently shown that retro-inverso isomerization decreases affinity toward both MDM2 and MDMX and that D-peptide ligands of MDM2 adopted left-handed helical conformations in both free and bound states. These later findings suggested a limited efficacy of retro-inverso strategy in functional and molecular mimicry of natural helical peptides [49].

1.4.2 Bcl-2 Proteins

The family of B-cell lymphoma 2 (Bcl-2) proteins comprises members with opposite functions, some responsible for cell survival, others for cell apoptosis, modulating the balance between cell survival and cell death [50]. The class of survival proteins includes anti-apoptotic members such as Bcl-2, Bcl-xL, Bcl-w, Mcl-1, and A1. The class of apoptosis proteins includes pro-apoptotic members such as Bax, Bak, and proteins of the BH3-only family. All members of Bcl-2 protein family are structurally related, containing at least one of four conserved Bcl-2 homology (BH) motifs, namely BH1, BH2, BH3, and BH4. In more detail, anti-apoptotic proteins contain all four BH domains, while pro-apoptotic proteins can conserve BH3 and other BH domains (e.g., Bax and Bak) or the BH3 domain only (this sub-family includes, among others, Bad, Bid, Bik, and Bim). Anti-apoptotic proteins promote cellular survival by trapping into a hydrophobic binding groove the helical BH3 domain of pro-apoptotic Bcl-2 family members, thus inhibiting their ability to induce cell death. If the pro-apoptotic proteins of the BH3-only family (which are upregulated in response to various cellular stresses, such as DNA damage or growth factor deprivation) [51] bind the anti-apoptotic members, Bax and Bak initiate the apoptotic process [50, 52].

In many tumor cells, Bcl-2 survival proteins are upregulated, causing a decrease in the apoptosis phenomenon, promoting survival and proliferation of the cells and their resistance to standard therapy.

Inhibition of interaction between anti-apoptotic Bcl-2 proteins and BH3 α-helix domains of pro-apoptotic proteins Bax and Bak was achieved by different compounds. Such compounds act as mimetics of the BH3 α-helix and inhibit survival proteins through binding at their BH3-binding pockets. Given the subtle differences among BH3-binding sites, Bcl-2 inhibitors usually target three or more anti-apoptotic proteins. The acylsulfonamide series of inhibitors from Abbott Laboratories had in ABT-737 the first inhibitor of the Bcl-2 family (Fig. 1.14) [53]. The compound is a BH3 mimetic and was reported to inhibit Bcl-xL, Bcl-2, and Bcl-w with an IC_{50} in the nanomolar range. Moreover, it showed activity in lymphoma and small cell lung carcinoma cell lines as well as primary patient-derived chronic lymphocytic leukemia cells. The low bioavailability of ABT-737 gave origin to further optimization. ABT-263 (Navitoclax) was thus identified and is currently undergoing clinical evaluation [54].

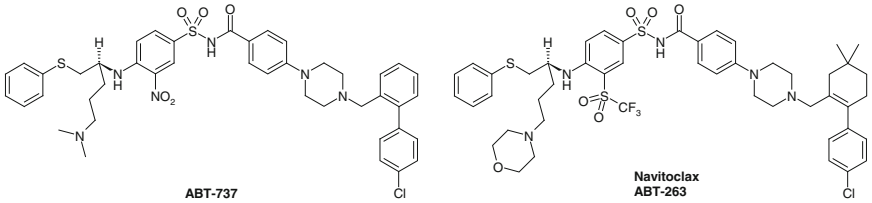
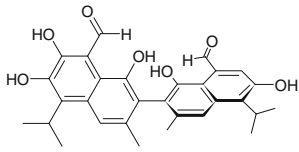


Fig. 1.14 Abbott's Bcl-2 family inhibitors

Fig. 1.15 Bcl-2, Bcl-xL, and Mcl-1 inhibitor



Another potent inhibitor of the Bcl-2 family members Bcl-2, Bcl-xL, and Mcl-1, is the natural product (R)-(-)-Gossypol (AT-101; Fig. 1.15). Similarly to ABT compounds, Gossypol mimics the BH3 domain of Bcl-2 pro-apoptotic proteins, thus interfering with the function of survival Bcl-2 proteins. Unfortunately, it has been reported that Gossypol failed to show efficacy in a phase II study on chemosensitive SCLC patients [55].

In the search for selective inhibitors of anti-apoptotic protein Mcl-1, a series of stabilized α-helix of Bcl-2 domains (SAHBs) has been designed based on the BH3 domains of pro- and anti-apoptotic Bcl-2 family members. Two cross-linking non-natural amino acids were substituted at the i, i + 4 positions of the non-interacting helical surface and stapled by ruthenium-catalyzed olefin metathesis. Interestingly, screening of the peptide library led to the identification of Mcl-1 BH3 helix itself as a potent and selective Mcl-1 inhibitor. Mcl-1 SAHB was shown to directly target Mcl-1, neutralize its interaction with pro-apoptotic Bak, and sensitize cancer cells to apoptosis [56]. X-ray crystallography confirmed that Mcl-1 SAHB is an α-helix interacting with MCL-1 at the BH3-binding groove of the anti-apoptotic protein (Fig. 1.16).

1.4.3 XIAP–Smac/DIABLO

X-linked inhibitor of apoptosis protein (XIAP) is an anti-apoptotic protein over-expressed in many human tumors that contributes to cancer cell resistance to chemotherapy. The protein contains three baculovirus IAP repeat called BIR domains, in which the linker sequence preceding the BIR2 domain is involved in the inhibition of both caspase-3 and caspase-7, whereas the BIR3 domain binds to

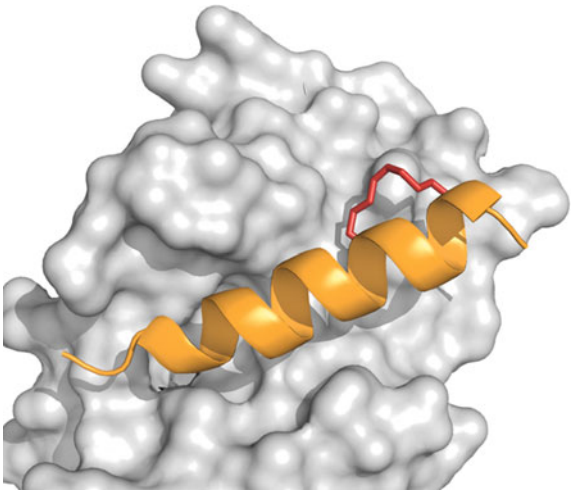


Fig. 1.16 X-ray structure of the complex between Mcl-1 SAHB peptide and Mcl-1. The hydrocarbon staple is shown as *red sticks*

and thereby inhibits caspase-9. The naturally occurring protein Smac/DIABLO (second mitochondria-derived activator of caspases/direct IAP-binding protein with low pI) is a tetrapeptide Ala-Val-Pro-Ile (AVPI) acting as XIAP inhibitor. Smac/DIABLO works against the anti-apoptotic function of XIAP by targeting both the BIR2 and the BIR3 domains of XIAP.

Small molecules able to reactivate apoptotic pathways blocked by aberrant XIAP activity have been obtained starting from the structure of AVPI peptide. SM-122 and SM-157, which have incorporated the central VP motif of AVPI, demonstrated to bind both the BIR2 and the BIR3 domains separately with high affinity, whereas SM-164 demonstrated to bind simultaneously to both domains (Fig. 1.17). The bivalent molecule displayed good affinity for XIAP in an in vitro competition assay and inhibited cellular proliferation inducing apoptosis of human leukemia cells [57, 58].

1.4.4 PDZ Domains

PDZ domains (an acronym combining the first letters of three proteins discovered to share this domain: Post-synaptic density protein, Drosophila disc large tumor suppressor, and Zonula occludens-1 protein) are widespread protein–protein interaction modules involved in many signaling systems relevant to development and maintenance of both pre- and post-synaptic structures and thus are subject to therapeutic intervention [59]. Compounds 3,289–8,625 and FJ9 are inhibitors of the PDZ domain of the Disheveled (Dvl) family of proteins (Fig. 1.18). Such

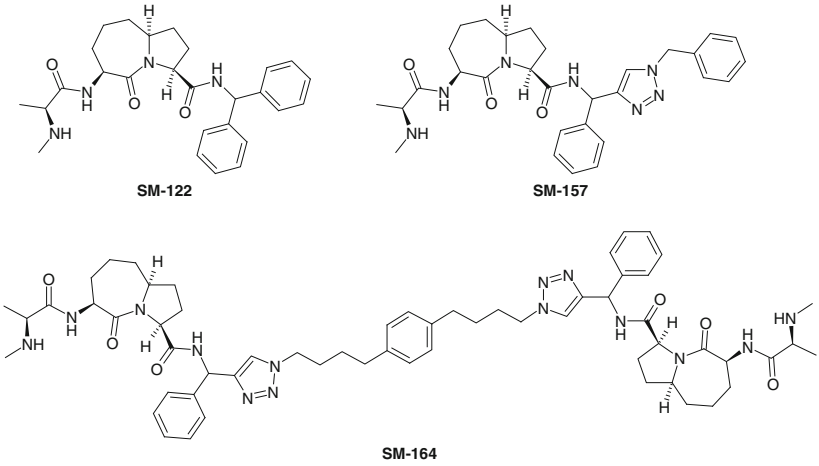


Fig. 1.17 Small molecules blocking aberrant XIAP activity

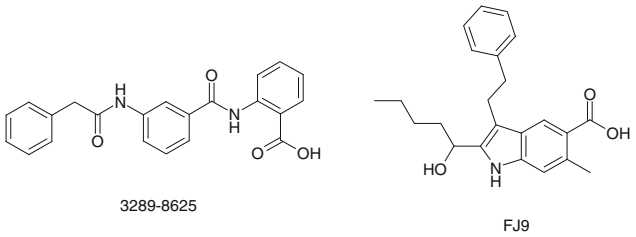


Fig. 1.18 Inhibitors of Disheveled PDZ domain

proteins are involved in Wnt signaling pathway and act directly downstream at Wnt receptors of the Frizzled family. In more detail, 3,289–8,625 compounds have been reported to interact with the PDZ domain of Dvl with micromolar affinity, block Wnt signaling, and reduce the growth rate of prostate cancer cell line PC-3 [60]. Similarly, FJ9 is able to downregulate Wnt signaling pathway and inhibit tumor growth in a mouse xenograft model by disrupting the interaction between the PDZ domain of Dvl and the Frizzled-7 Wnt receptor [61].

1.4.5 HIF-1 Pathway

Hypoxia-inducible factor 1 (HIF-1) is a heterodimeric transcription factor composed of an α and β subunit. The HIF-1 complex mediates cellular response to hypoxic conditions and it is found to be overexpressed in many forms of cancer. Its interaction with the transcription coactivator CBP/p300 represents a key event in the signal transduction pathway, and it is essential for survival of hypoxic cells.

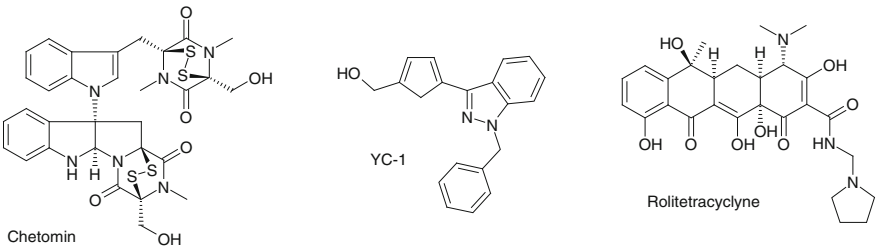


Fig. 1.19 Inhibitors of HIF-1 PPIs

Blocking the HIF-1 activity in tumor cells by inhibiting HIF-1 PPIs is considered a drug discovery opportunity. Accordingly, small molecules chetomin and YC-1 (Fig. 1.19), as well as a polypeptide, have been found to inhibit the binding between HIF-1 α and CBP/p300, while rolitetracycline was shown to block the HIF-1 α/β dimerization [62]. However, despite the micromolar activity in cell-free assays, rolitetracycline failed to inhibit HIF-1-dependent luciferase expression in a cell-based assay and also did not affect HIF-1 α/β interaction either in intact cells or in nuclear proteins from hypoxia-treated cells. Failure of activity in cell-based assays might be ascribed to either poor membrane permeability or low solubility, or both [63].

1.4.6 Tubulin Polymerization

Other examples of the use of natural products to modulate PPIs can be found among inhibitors of either tubulin polymerization or depolymerization. Microtubules consist of α/β-tubulin heterodimers and are formed during cell division. Vinca alkaloids, for example, vincristine and vinblastine, and colchicines are inhibitors of tubulin polymerization. Inhibitors of depolymerization are instead taxanes (diterpenes produced by the plants of the genus Taxus, e.g., paclitaxel) and epothilones (Fig. 1.20).

These two different functional classes regulate allosterically the tubulin oligomerization process binding to different regions of the α/β-tubulin heterodimer [64]. There are also known small molecule inhibitors of tubulin, such as HTI-286, a peptidomimetic synthetic analog of the marine sponge targeting prostate cancer [65], and BPR-0L-075, inhibitor of the tubulin polymerization through the binding with colchicine-binding site of tubulin [66].

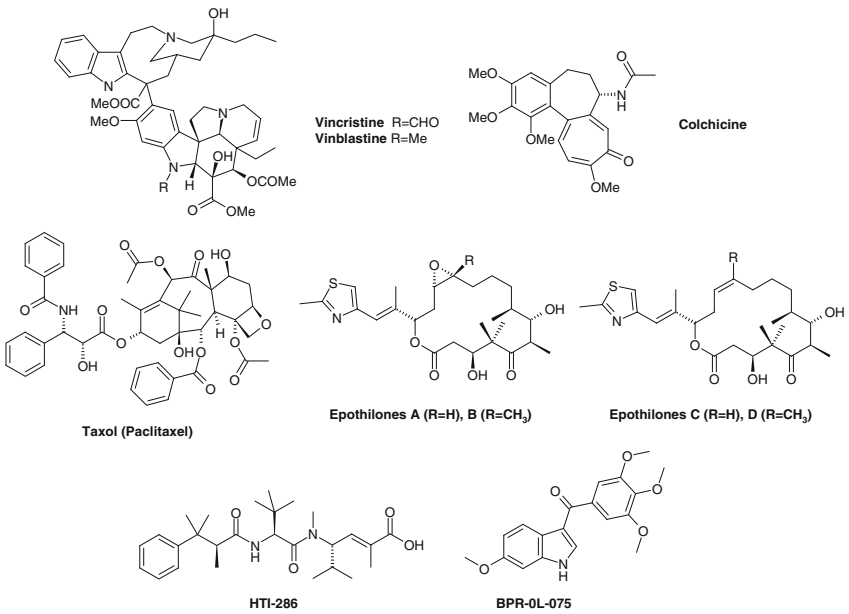
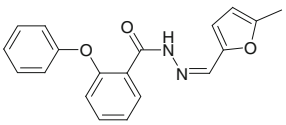


Fig. 1.20 Inhibitors of tubulin polymerization/depolymerization

1.4.7 β-Catenin Transcription Factors

β-catenin is a multifunctional protein, and binding to its downstream signaling partners (Lymphoid enhancer factor-1, LEF-1, and T cell factors, Tcf-1, Tcf-3, and Tcf-4) in the Wnt signal transduction pathway activates the transcription of target genes involved in cell cycle regulation. Increased levels of β-catenin have been associated with several human cancers [67]. In tumor cells, β-catenin is stabilized, resulting in the stimulation of LEF/Tcf transcription factors and subsequent cellular proliferation. The complexes between β-catenin and Tcf family members show a very broad and extended binding region, difficult to be targeted by small molecules; furthermore considering that β-catenin is involved in protein complex formations with other downstream partners such as APC and axin, which have an important physiological role in tumor growth suppression, it is also necessary to have selectivity to avoid the inhibition of these interactions.

Compound PNU74654 (Fig. 1.21) was shown to inhibit the Wnt signaling pathway by blocking the interaction between β-catenin and Tcf-4 [68]. PNU74654 was identified by docking of 17,700 compounds of the Pharmacia corporate to defined hot spots present in the Tcf3/Tcf4-binding surface of β-catenin. 22 compounds were selected for biophysical screening (NMR and ITC), which revealed three compounds that were able to specifically bind to β-catenin and to compete with Tcf4. Among them, PNU74654 was shown to bind with the greatest affinity ($K_d = 450$ nM).



PNU74654

Fig. 1.21 Inhibitor of the interaction between β-catenin and Tcf-4 transcription factor

1.4.8 NOTCH Proteins

NOTCH proteins (NOTCH1–4 in humans) are single-pass transmembrane receptors that regulate cellular differentiation, proliferation, and death. The alteration of NOTCH1 signaling pathway is directly implicated in the pathogenesis of several disease states, including T cell acute lymphoblastic leukemia (T-ALL) [69]. Binding of ligand proteins (Jagged-1 and Jagged-2 and Delta-like-1, Delta-like-2, and Delta-like-4) to the extracellular domain of NOTCH1 initiates proteolytic processes catalyzed by both ADAM family metalloprotease and γ-secretase complex (Fig. 1.22). The released intracellular domain of NOTCH1 (NICD) translocates to the nucleus and interacts with the DNA-bound transcription factor CSL. The complex NICD–CSL subsequently interacts with co-activator proteins of the mastermind-like (MAML) family, and the resulting ternary complex is responsible for the activation of NOTCH-dependent target genes.

A fragment of MAML1 (residues 13–74) has been shown to antagonize NOTCH signaling and cell proliferation when expressed in T-ALL cell lines. This polypeptide forms a nearly continuous α-helix, suggesting that the NOTCH transactivation complex might be suitable for targeting by helix mimetics such as hydrocarbon-stapled α-helical peptides. Accordingly, the structure of human NOTCH1 ternary complex (Fig. 1.23), and in particular a 16-amino acid stretch of MAML1 targeting both NICD and CSL, was used as the basis for the design of a series of stapled α-helical peptides (SAHMs). Among them, peptide labeled as SAHM1 was shown to bind NOTCH1 transactivation complex with high affinity

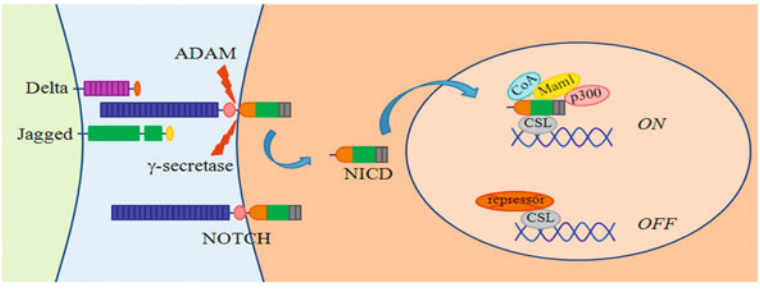


Fig. 1.22 NOTCH signaling pathway

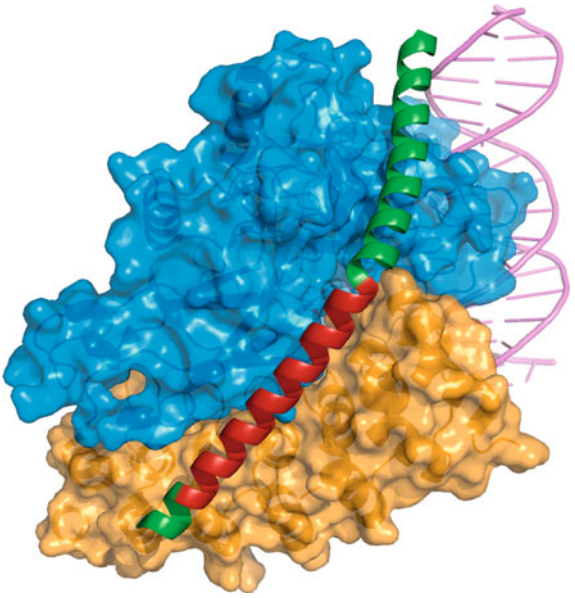


Fig. 1.23 X-ray structure of NOTCH1 ternary complex. MAML1 (*green*), CSL (*blue*), NICD (*Orange*), and DNA (*pink*). The portion of MAML-1 binding at the CSL–NICD interface is highlighted in *red*

and competitively with MAML1, and to specifically repress NOTCH1 target gene expression [70].

1.4.9 HIV

Another interesting illustration comes from the capsid domain of the human immunodeficiency virus type 1 (HIV-1) Gag polyprotein, which is a critical determinant of virus assembly and for this reason a potential target for developing drugs for AIDS therapy. A 12-mer α-helical peptide (CAI) [71] was reported to disrupt immature- and mature-like capsid particle assembly in vitro, but it failed to inhibit HIV-1 in cell culture for low permeability. Moreover, the X-ray crystal structure of CAI in complex with the C-terminal domain of capsid was determined [72]. Starting from these structural data, Debnath and coworkers designed a stapled peptide with the stabilized α-helical structure of CAI and converted it to a cell-penetrating peptide (CPP), labeled as NYAD-1 [73]. NYAD-1 showed enhanced α-helicity and was able to inhibit a large panel of HIV-1 isolates in cell culture with low micromolar potency. Subsequently, the structure of the complex between the C-terminal domain of HIV-1 capsid and NYAD-13, a soluble peptide analog of NYAD-1, was determined by NMR (Fig. 1.24) [74].

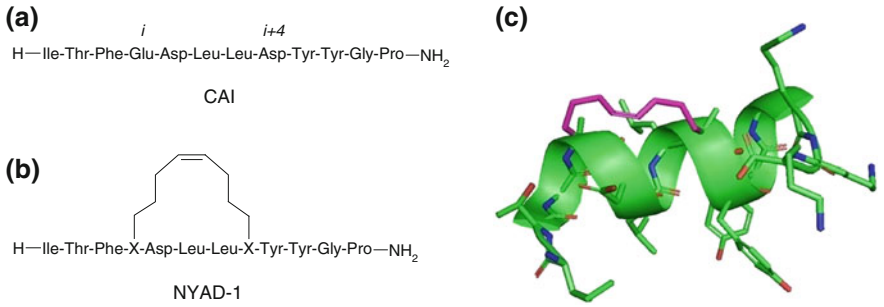


Fig. 1.24 (**a**) Sequence of CAI peptide. (**b**) Schematic representation of the stapled NYAD-1 peptide. (**c**) Structure of NYAD-13 peptide (the hydrocarbon staple is shown as *magenta sticks*), extracted from the NMR structure of the complex between the stapled peptide and the C-terminal domain of HIV-1 capsid protein

1.4.10 S100B-p53 Interaction

S100B is a small calcium-binding protein member of the highly conserved S100 family, which includes 22 known members implicated in intracellular and extracellular regulatory activities. S100B interacts with the tumor suppressor protein p53 altering its function, and it is considered an interesting pharmaceutical target due its overexpression in several tumor cells, particularly in metastatic melanoma. A few small molecules that bind human S100B have been identified by Weber and coworkers using an NMR-based screening [75–77], while Arendt et al. identified inhibitors of the highly homologous bovine protein with a similar methodology [78].

Furthermore, an in silico fragment screening targeting the S100B–p53-binding interface was performed taking into consideration an ensemble of S100B conformations derived from available NMR structures. 280 fragments were selected for biophysical evaluation, and five of them (corresponding to three different scaffolds) were confirmed as selective binders to the peptide-binding site, with an estimated K_d value in the range of 0.1–1.4 mm and ligand efficiency in the range of 0.19–0.26 kcalmol^{-1} [79].

1.4.11 Eph-ephrin Signaling

Eph receptors are a family of tyrosine kinase receptors involved in a number of cellular processes, including migration, adhesion, proliferation, survival, and differentiation. The extracellular domain of Eph receptors interacts with cell surface-associated ephrin proteins on neighboring cells, resulting in the activation of several downstream signals [80]. Altered expression or deregulated function of Eph receptors and/or ephrins constitutes the basis for the central role played by the

Eph-ephrin system in a large variety of human cancers [81]. For example, EphA2 upregulation was correlated with tumor stage and progression, and the expression of EphA2 in non-transformed cells induced malignant transformation and conferred tumorigenic potential. These findings made the identification of small molecule inhibitors of the EphA2–ephrinA1 interaction an important therapeutic opportunity [82]. Similarly, EphB4 is a member of this family which has both tumor-suppressing and tumor-promoting activities in breast cancer. Understanding the role of EphB4 in tumorigenesis may allow the development of new anticancer therapies [83].

1.5 Where are We? Have Expectations been Met?

Historical PPI inhibition drug discovery projects, such as p53–MDM2 or Bcl family inhibitors, have recently entered clinical evaluation. In the next few years, it will be possible to ascertain relevance of interaction inhibition versus drug efficacy and toxicities. Moreover, successful examples will shed the light on best drug design and medicinal chemistry avenues to follow to deliver clinical candidates. Current examples indicate that PPI inhibitor candidates are characterized by physico-chemical properties outside standard drug-like properties, in particular in terms of size and ligand efficiency. Medicinal chemistry expertise has taken up the challenge and has been instrumental in addressing emerging ADMET issues.

1.6 Perspective

In the last few years, there has been a major focus in the high-throughput application of experimental technologies to identify and characterize PPI, assisted by in silico predictions approaches and PPI databases creation. The next few years will see a boost in the development of biochemical and cellular high-throughput screening technologies to identify suitable chemical entry points for lead optimization programs. Computational chemistry approaches, such as protein–protein docking and molecular dynamics simulations supported by reduced CPU time, are expected to allow rational drug design. In particular, it may be possible to run in silico virtual screening campaigns delivering potential hits targeting meta-states aimed at freezing protein conformational rearrangements upon protein–protein contacts formation. It may also be possible to see examples of small molecules, acting as chaperones, capable of stabilizing protein complexes. In terms of library design, it will be necessary to expand and explore beyond the classical drug-like boundaries. In particular, in order to identify chemical tools useful to modulate classical PPI, size does matter. Thus, there might be a resurgence of focused libraries of peptoids or large small molecules.

References

1. Keskin O, Gursoy A, Ma B, Nussinov R (2008) Principles of protein–protein interactions: what are the preferred ways for proteins to interact? Chem Rev 108:1225–1244
2. De S, Krishnadev O, Srinivasan N, Rekha N (2005) Interaction preferences across protein–protein interfaces of obligatory and non-obligatory components are different. BMC Struct Biol 5:15
3. Tsai C, Ma B, Nussinov R (2009) Protein-protein interaction networks: how can a hub protein bind so many different partners? Trends Biochem Sci 34(12):594–600
4. Zinzalla G, Thurston DE (2009) Targeting protein–protein interactions for therapeutic intervention: a challenge for the future. Future Med Chem 1:65–93
5. Sperandio O, Reynès CH, Camproux AC, Villoutreix BO (2010) Rationalizing the chemical space of protein–protein interaction inhibitors. Drug Discovery Today 15(5/6):220–229
6. Chène P (2006) Drugs targeting protein–protein interactions. ChemMedChem 1:400–411
7. Hajduk PJ, Huth JR, Fesik SW (2005) Druggability indices for protein targets derived from NMR-based screening data. J Med Chem 48:2518–2525
8. Zhao N, Pang B, Shyu C, Korkin D (2011) Structural similarity and classification of protein interaction interfaces. PlosOne, 6, e19554:1–14
9. Fuller JC, Burgoyne NJ, Jackson RM (2009) Predicting druggable binding sites at the protein–protein interface. Drug Discovery Today 14(3/4):155–161
10. Metz A, Pfleger C, Kopitz H, Pfeiffer-Marek S, Baringhaus KH, Gohlke H (2012) Hot spots and transient pockets: predicting the determinants of small-molecule binding to a protein–protein interface. J Chem Inf Model 52:120–133
11. Arkin MR, Wells JA (2004) Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. Nat Rev Drug Discovery 3:301–317
12. Petschnigg J, Snider J, Stagljar I (2011) Interactive proteomics research technologies: recent applications and advances. Curr Opin Biotechnol 22:50–58
13. Macchiarulo A, Giacchè N, Carotti A, Baroni M, Cruciani G, Pellicciari R (2008) Targeting the conformational transitions of MDM2 and MDMX: insights into dissimilarities and similarities of p53 recognition. J Chem Inf Model 48:1999–2009
14. Hajduk PJ, Huth JR, Tse C (2005) Predicting Protein Druggability. Drug Discovery Today 10:1675–1682
15. Diller DJ, Hobbs DW (2004) Deriving knowledge through data mining high-throughput screening data. J Med Chem 47:6373–6383
16. Zapatero CA (2007) Ligand efficiency indices for effective drug discovery. Expert Opin Drug Discovery 2:469–488
17. Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. Nature 450:1001–1009
18. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26
19. Mason JM (2010) Design and development of peptides and peptide mimetics as antagonists for therapeutic intervention. Future Med Chem 2:1813–1822
20. Rubinstein M, Niv MY (2009) Peptidic modulators of protein–protein interactions: progress and challenges in computational design. Biopolymers 91:505–513
21. Ullman CG, Frigotto L, Cooley RN (2011) In vitro methods for peptide display and their applications. Briefings Funct Genom 10:125–134
22. Kessler H (1982) Conformation and biological activity of cyclic peptides. Angew Chem Int Ed Engl 21(7):512–523
23. Henchey LK, Jochim AL, Arora PS (2008) Contemporary strategies for the stabilization of peptides in the a-helical conformation. Curr Opin Chem Biol 12:692–697
24. Yin H, Hamilton AD (2005) Strategies for targeting protein–protein interactions. Angew Chem Int Ed 44:4130–4163

25. Blazer LL, Neubig RR (2009) Small molecule protein–protein interaction inhibitors as CNS therapeutic agents: current progress and future hurdles. Neuropsychopharmacology 34:126–141
26. Fry DC, Vassilev LT (2005) Targeting protein–protein interactions for cancer therapy. J Mol Med 83:955–963
27. Aisen PS, Saumier D, Briand R, Laurin J, Gervais F, Tremblay P, Garceau D (2006) A phase II study targeting amyloid-beta with 3APS in mild-to-moderate Alzheimer disease. Neurology 67:1757–1763
28. Rafii MS, Aisen PS (2009) Recent developments in Alzheimers disease therapeutics. BMC Medicine 7:7
29. Bolognesi ML, Cavalli A, Melchiorre C (2009) Memoquin: A multi-target–directed ligand as an innovative therapeutic opportunity for Alzheimers disease. Neurotherapeutics 6:152–162
30. Gestwicki JE, Crabtree GR, Graef IA (2004) Harnessing chaperones to generate small-molecule inhibitors of amyloid beta aggregation. Science 306:865–869
31. Ryan DP, Matthews JM (2005) Protein–protein interactions in human disease. Curr Opin Struct Biol 15:441–446
32. Neubig RR, Siderovski DP (2002) Regulators of G-protein signalling as new central nervous system drug targets. Nat Rev Drug Discovery 1:187–197
33. Popowicz GM, Dömling A, Holak TA (2011) The structure-based design of Mdm2/MDMX–p53 inhibitors gets serious. Angew Chem Int Ed 50:2680–2688
34. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science 274(5289):948–953
35. Popowicz GM, Czarna A, Rothweiler U, Szwagierczak A, Krajewski M, Weber L, Holak TA (2007) Molecular basis for the inhibition of p53 by Mdmx. Cell Cycle 6:2386–2392
36. Vassilev LT (2005) p53 activation by small molecules: application in oncology. J Med Chem 48:4491–4499
37. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303:844–848
38. Ohnstad HO, Paulsen EB, Noordhuis P, Berg M, Lothe RA, Vassilev LT, Myklebost O (2011) MDM2 antagonist Nutlin-3a potentiates antitumour activity of cytotoxic drugs in sarcoma cell lines. BMC Cancer 11:211–239
39. Brennan RC, Federico S, Bradley C, Zhang J, Flores-Otero J, Wilson M, Stewart CF, Zhu F, Guy K, Dyer MA (2011) Targeting the p53 pathway in Retinoblastoma with subconjunctival Nutlin-3a. Cancer Res 71:4205–4213
40. Tabe Y, Sebasigari D, Jin L, Rudelius M, Davies-Hill T, Miyake K, Miida T, Pittaluga S, Raffeld M (2009) MDM2 antagonist nutlin-3 displays antiproliferative and proapoptotic activity in mantle cell lymphoma. Clin Cancer Res 15:933–942
41. Yu S, Qin D, Shangary S, Chen J, Wang G, Ding K, McEachern D, Qiu S, Nikolovska-Coleska Z, Miller R, Kang S, Yang D, Wang S (2009) Potent and orally active small-molecule inhibitors of the MDM2-p53 interaction. J Med Chem 52:7970–7973
42. Grasberger BL, Lu T, Schubert C, Parks DJ, Carver TE, Koblish HK, Cummings MD, LaFrance LV, Milkiewicz KL, Calvo RR, Maguire D, Lattanze J, Franks CF, Zhao S, Ramachandren K, Bylebyl GR, Zhang M, Manthey CL, Petrella EC, Pantoliano MW, Deckman IC, Spurlino JC, Maroney AC, Tomczuk BE, Molloy CJ, Bone RF (2005) Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in Cells. J Med Chem 48:909–912
43. Hardcastle IR, Liu J, Valeur E, Watson A, Ahmed SU, Blackburn TJ, Bennaceur K, Clegg W, Drummond C, Endicott JA, Golding BT, Griffin RJ, Gruber J, Haggerty K, Harrington RW, Hutton C, Kemp S, Lu X, McDonnell JM, Newell DR, Noble MEM, Payne SL, Revill CH, Riedinger C, Xu Q, Lunec J (2011) Isoindolinone inhibitors of the murine double minute 2 (MDM2)-p53 protein–protein interaction: structure-activity studies leading to improved potency. J Med Chem 54:1233–1243

44. Riedinger C, Endicott JA, Kemp SJ, Smyth LA, Watson A, Valeur E, Golding BT, Griffin RJ, Hardcastle IR, Noble ME, McDonnell JM (2008) Analysis of chemical shift changes reveals the binding modes of isoindolinone inhibitors of the MDM2-p53 interaction. J Am Chem Soc 130:16038–16044
45. Popowicz GM, Czarna A, Wolf S, Wang K, Wang W, Dömling A, Holak TA (2010) Structures of low molecular weight inhibitors bound to MDMX and MDM2 reveal new approaches for p53-MDMX/MDM2 antagonist drug discovery. Cell Cycle 9:1104–1111
46. Czarna A, Popowicz GM, Pecak A, Wolf S, Dubin G, Holak TA (2009) High affinity interaction of the p53 peptide-analogue with human Mdm2 and MDMX. Cell Cycle 8:1176–1184
47. Bernal F, Tyler AF, Korsmeyer SJ, Walensky LD, Verdine GL (2007) Reactivation of the p53 tumor suppressor pathway by a stapled p53 peptide. J Am Chem Soc 129:2456–2457
48. Sakurai K, Chung HS, Kahne D (2004) Use of a retroinverso p53 peptide as an inhibitor of MDM2. J Am Chem Soc 126:16288–16289
49. Li C, Pazgier M, Li J, Li C, Liu M, Zou G, Li Z, Chen J, Tarasov SG, Lu WY, Lu W (2010) Limitations of peptide retro-inverso isomerization in molecular mimicry. J Biol Chem 285:19572–19581
50. Sleebs BE, Czabotar PE, Fairbrother WJ, Fairlie WD, Flygare JA, Huang DCS, Kersten WJA, Koehler MFT, Lessene G, Lowes K, Parisot JP, Smith BJ, Smith ML, Souers AJ, Street IP, Yang H, Baell JB (2011) Quinazoline Sulfonamides as dual binders of the proteins B-Cell Lymphoma 2 and B-Cell Lymphoma extra long with potent proapoptotic cell-based activity. J Med Chem 54:1914–1926
51. Hinds MG, Smits C, Fredericks-Short R, Risk JM, Bailey M, Huang DCS, Day CL (2007) Bim, Bad and Bmf: intrinsically unstructured BH3-only proteins that undergo a localized conformational change upon binding to prosurvival Bcl-2 targets. Cell Death Differ 14:128–136
52. Reed JC, Zha H, Aime-Sempe C, Takayama S, Wang HG (1996) Structure-function analysis of Bcl-2 family proteins, Regulators of programmed cell death. Adv Exp Med Biol 406:99–112
53. Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, Rosenberg SH (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature 435:677–681
54. Gandhi L, Camidge DR, de Oliveira MR, Bonomi P, Gandara D, Khaira D, Hann CL, McKeegan EM, Litvinovich E, Hemken PM, Dive C, Enschede SH, Nolan C, Chiu Y, Busman T, Xiong H, Krivoshik AP, Humerickhouse R, Shapiro GI, Rudin CM (2011) Phase I study of navitoclax (ABT-263), a novel Bcl-2 family inhibitor, in patients with small-cell lung cancer and other solid tumors. J Clin Oncol 29:909–916
55. Baggstrom MQ, Qi Y, Koczywas M, Argiris A, Johnson EA, Millward MJ, Murphy SC, Erlichman C, Rudin CM, Govindan R (2011) Mayo Phase 2 Consortium, California Consortium: A phase II study of AT-101 (Gossypol) in chemotherapy-sensitive recurrent extensive-stage small cell lung cancer. J Thoracic Oncol 6:1757–1760
56. Stewart ML, Fire E, Keating AE, Walensky LD (2010) The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. Nat Chem Biol 6:595–601
57. Berg T (2008) Small-molecule inhibitors of protein–protein interactions. Curr Opin Drug Discov Dev 11:666–674
58. Rajapakse H (2007) Small molecule inhibitors of the XIAP protein–protein interaction. Curr Top Med Chem 7:966–971
59. Houslay MD (2009) Disrupting specific PDZ domain-mediated interactions for therapeutic benefit. Br J Pharmacol 158:483–485

60. Grandy D, Shan J, Zhang X, Rao S, Akunuru S, Li H, Zhang Y, Alpatov I, Zhang XA, Lang RA, Shi DL, Zheng JJ (2009) Discovery and characterization of a small molecule inhibitor of the PDZ domain of dishevelled. J Biol Chem 284:16256–16263
61. Fujii N, You L, Xu Z, Uematsu K, Shan J, He B, Mikami I, Edmondson LR, Neale G, Zheng J, Guy RK, Jablons DM (2007) An antagonist of dishevelled protein–protein interaction suppresses β-Catenin–dependent tumor cell growth. Cancer Res 67:573–579
62. Poon E, Harrisa AL, Ashcroft M (2009) Targeting the hypoxia-inducible factor (HIF) pathway in cancer. Expert Rev Mol Med 11:e26
63. Park EJ, Kong D, Fisher R, Cardellina J, Shoemaker RH, Melillo G (2006) Targeting the PAS-A domain of HIF-1a for development of small molecule inhibitors of HIF-1. Cell Cycle 5(16):1847–1853
64. Arkin M (2005) Protein–protein interactions and cancer: small molecules going in for the kill. Curr Opin Chem Biol 9:317–324
65. Hadaschik BA (2008) Targeting prostate cancer with HTI-286, a synthetic analog of the marine sponge product hemiasterlin. Int J Cancer 122:2368–2376
66. Kuo CC, Hsieh HP, Pan WY, Chen CP, Liou JP, Lee SJ, Chang YL, Chen LT, Chen CT, Chang JY (2004) BPR0L075, a novel synthetic indole compound with antimitotic activity in human cancer cells, exerts effective antitumoral activity in vivo. Cancer Res 64:4621–4628
67. Takayama T, Shiozaki H, Shibamoto S, Oka H, Kimura Y, Tamura S, Inoue M, Monden T, Ito F, Monden M (1996) Beta-catenin expression in human cancers. Am J Pathol 148:39–46
68. Trosset J, Dalvit C, Knapp S, Fasolini M, Veronesi M, Mantegani S, Gianellini LM, Catana C, Sundström M, Stouten PFW, Moll JK (2006) Inhibition of protein–protein interactions: the discovery of druglike β-catenin inhibitors by combining virtual and biophysical screening. Proteins: Struct Funct Bioinf 64:60–67
69. Weng AP, Ferrando AA, Lee W, Morris JP, Silverman LB, Sanchez-Irizarry C, Blacklow SC, Look AT, Aster JC (2004) Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science 306:269–271
70. Moellering RE, Cornejo M, Davis TN, Del Bianco C, Aster JC, Blacklow SC, Kung AL, Gilliland DG, Verdine GL, Bradner JE (2009) Direct inhibition of the NOTCH transcription factor complex. Nature 462:182–188
71. Sticht J, Humbert M, Findlow S, Bodem J, Müller B, Dietrich U, Werner J, Kräusslich HG (2005) A peptide inhibitor of HIV-1 assembly in vitro. Nat Struct Mol Biol 12:671–677
72. Ternois F, Sticht J, Duquerroy S, Krausslich HG, Rey FA (2005) The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. Nat Struct Mol Biol 12:678–682
73. Zhang H, Zhao Q, Bhattacharya S, Waheed AA, Tong X, Hong A, Heck S, Curreli F, Goger M, Cowburn D, Freed EO, Debnath AK (2008) A Cell-penetrating helical peptide as a potential HIV-1 inhibitor. J Mol Biol 378:565–580
74. Bhattacharya S, Zhang H, Debnath AK, Cowburn D (2008) Solution structure of a hydrocarbon stapled peptide inhibitor in complex with monomeric C-terminal domain of HIV-1 capsid. J Biol Chem 283:16274–16278
75. Markowitz J, Chen I, Gitti R, Baldisseri DM, Pan Y, Udan R, Carrier F, MacKerell AD, Weber DJ (2004) Identification and characterization of small molecule inhibitors of the calcium-dependent S100B–p53 tumor suppressor interaction. J Med Chem 47:5085–5093
76. Markowitz J, MacKerell AD, Weber DJ (2007) A Search for inhibitors of S100B, a member of the S100 family of calcium-binding proteins. Mini-Rev Med Chem 7:609–616
77. Charpentier TH, Wilder PT, Liriano MA, Varney KM, Zhong S, Coop A, Pozharski E, MacKerell AD, Toth EA, Weber DJ (2009) Small molecules bound to unique sites in the target protein binding cleft of calcium-bound S100B as characterized by nuclear magnetic resonance and X-ray crystallography. Biochemistry 48:6202–6212
78. Arendt Y, Bhaumik A, Del Conte R, Luchinat C, Mori M, Porcu M (2007) Fragment docking to S100 proteins reveals a wide diversity of weak interaction sites. ChemMedChem 2:1648–1654

79. Agamennone M, Cesari L, Lalli D, Turlizzi E, Del Conte R, Turano P, Mangani S, Padova A (2010) Fragmenting the S100B–p53 interaction: combined virtual/biophysical screening approaches to identify ligands. ChemMedChem 5:428–435
80. Pasquale EB (2008) Eph-ephrin bidirectional signaling in physiology and disease. Cell 133:38–52
81. Ireton RC, Chen J (2005) EphA2 receptor tyrosine kinase as a promising target for cancer therapeutics. Curr Cancer Drug Targets 5:149–157
82. Giorgio C, Mohamed IH, Flammini L, Barocelli E, Incerti M, Lodola A, Tognolini M (2011) Lithocholic acid is an Eph-ephrin ligand interfering with Eph-kinase activation. PLoS One, 6, e18128
83. Noren NK, Pasquale EB (2007) Paradoxes of the EphB4 receptor in cancer. Cancer Res 67:3994–3997

Chapter 2
Protein–Protein Interaction Inhibitors: Case Studies on Small Molecules and Natural Compounds

Stefania Ferrari, Federica Pellati and Maria Paola Costi

Abbreviations

Aβ	β-amyloid
AβPP	β-amyloid protein precursor
Aib	α-aminoisobutyric acid
BIR	Baculoviral inhibitory repeat
c-kit	Tyrosine kinase receptor
dTMP	2′-deoxythymidine-5′-monophosphate
dUMP	2′-deoxyuridine-5′-monophosphate
FRET	Förster resonance energy transfer
GAPs	GTPase-activating proteins
GEFs	Guanine nucleotide exchange factors
HDM2	Human double minute 2
HeLa	Human cervical carcinoma cancer cell line
HTS	High-throughput screening
hTS	Human thymidylate synthase
IAPs	Inhibitors of apoptosis proteins
MAP	Mitogen-activated protein
MDM2	Mouse double minute 2
NCI	National Cancer Institute
NSAIDs	Non-steroidal anti-inflammatory drugs
O-PHDEs	Oxy-polyhalogenated diphenyl ethers
PAC	Proteasome assembling chaperone
PBD	Polo-box domain
PDPA	1,3-propanediphosphonic acid

S. Ferrari · F. Pellati · M. P. Costi (✉)
Department of Life Sciences, University of Modena and Reggio Emilia,
Via Campi 183, 41125 Modena, Italy
e-mail: mariapaola.costi@unimore.it

S. Ferrari
e-mail: stefania.ferrari@unimore.it

F. Pellati
e-mail: federica.pellati@unimore.it

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4_2,

PI3K	Phosphoinositol-3′-kinase
PPIs	Protein–protein interactions
PS	Presenilin
SCF	Stem cell factor
Tcf4	T-cell factor 4
XIAP	X-linked Inhibitor of apoptosis protein

2.1 Introduction

Many biological functions involve the formation of protein–protein complexes, and the inhibition of this process has garnered significant interest in pharmaceutical research investigating novel therapies for several human diseases. From an evolutionary perspective, proteins have evolved to optimize and differentiate their functions, a process that is mediated by the modulation of the interacting surfaces. Protein–protein interactions (PPIs) represent another dimension of the structural properties of proteins and are the essence of the interactome, the ensemble of all complexes generated through PPIs present at the cellular level. The aggregation states of a protein monomer, for example, can cover a wide range of complexes; the same protein in monomeric form may have a different function with respect to its aggregated form [1]. This phenomenon enlarges the biological space considerably and increases its complexity. PPIs can occur between two (or more) identical protein sequences (homo-oligomers) or between two different protein sequences (hetero-oligomers) [2]. PPIs can also be distinguished on the basis of the stability of the complex: there are permanent complexes (structurally obligate oligomers generally fall into this category) and transient complexes. In structurally and/or functionally obligate oligomers, single monomers cannot exist as a long-lived entity or they are inactive in vivo. The multiple functions of proteins may account for their high structural complexity. In the case of enzymes, their catalytic functions are mostly located in the active site region, while the other protein domains, both solvent accessible or non-solvent accessible, may be responsible for different functions. Many pathological processes are governed by PPIs; therefore, designing molecules that interfere with the formation of these complexes is one of the recent challenges in drug design [2].

So far, only a minimal fraction of the estimated 650,000 PPIs that comprise the human interactome are known, and only a few categories have been described [3]. These interactions control processes involved in both normal and pathological pathways, which include signal transduction, cell adhesion, cellular proliferation, growth, differentiation, viral self-assembly, programmed cell death, and cytoskeleton structure. PPIs that play a role in signal transduction or cell regulation are often temporary and weak [4, 5].

PPIs are tentatively organized as follows: (a) PPIs (high molecular mass), which include membrane receptors; (b) oligomeric proteins (monomer–monomer

interactions) that can be homo- or heteromeric; (c) peptide–protein interactions; and (d) protein–antibody interactions (immunoconjugates or immunocomplexes). In category (a), we include those proteins that often use the complex networks of interactions to produce sophisticated signaling networks that are capable of well-tuned and highly adaptive responses to environmental stimuli, such as programmed cell death [6–9].

Because signal transduction is related to how the extracellular environment exerts its effect on cellular status and function, many steps in signal transduction pathways involve proteins that represent potential drug targets [10]. Category (b) is mostly related to the self-assembly of proteins in different aggregates depending on the local cellular environment [11]. Category (c) includes all of the regulatory networks that are part of the peptidome (composed of all the small peptides that have often unknown regulatory functions in the cells) [12]. Category (d) includes antibody–antigen complexes, which are a well-described type of protein–protein interaction. This complex is characterized by highly specific molecular recognition. At the moment, this type of protein–protein interaction has not been studied from a drug discovery perspective but is mostly important from a functional point of view.

The experimental detection of protein complexes is very common. Biologists characterize new protein complexes to identify the total number of ligands that establish interactions with the target protein by co-precipitating the ligand and protein in immunocomplexes, using affinity tagging and pull-down assays or using two hybrid methods to understand their functions [13]. However, the mechanistic role of the identified complexes is not easy to establish. To this aim, unnatural amino acids have been inserted through molecular biology approaches using orthogonal acyl-t-RNA transferases to generate mutated proteins with altered functions [14]. The possibility of studying the biological role of PPIs in a native cellular environment is fundamental to avoid artifacts and drive data interpretations. X-ray crystallographers generate protein crystals that often show oligomeric structures. At the moment, crystallization experiments can only be performed at high protein concentrations, and therefore, protein aggregation depends on the equilibrium constants that govern the relative affinity of the monomers. Other biophysical methods are also important in efforts to understand the multimeric protein state. Among them, NMR can be used to determine structural information about the aggregation state of a protein. Although these experimental approaches are not performed in conditions that are fully physiologically relevant, they can be used to establish that the protein can self-assemble with a recognition code that is not easy to rationalize. So far, only a small number of complexes have been considered as druggable targets. Such intricate biological systems cannot be cost-effectively tackled using conventional high-throughput screening (HTS) methods. Some researchers argue that accomplishing this goal will be difficult, if not impossible, because the target binding site is a multipoint interface between two proteins with sufficient stabilizing force to bind two molecules together. Therefore, there is competition between the ligand and the partner protein. However, the development of PPIs inhibitors is becoming more and more feasible [15].

The increased enthusiasm for PPI inhibitors is not just restricted to preclinical programs, and companies are advancing a handful of such compounds through clinical trials [16].

2.2 Characterizing the Binding Site

Proteins interact in complicated ways because their shapes are so vastly complex. Amino acid side chains that extend from the body of the molecule create cavities or bumps of different shapes and sizes. Proteins maximally exploit this structural diversity, producing binding pockets and recognition sites with varying degrees of specificity. In general, a core interaction area and a limit interaction area are recognized in PPIs. Core interaction area involves the highest conserved residues, while the limit interaction area is less crucial for the interaction and is less conserved.

Most current drugs target the binding site of a protein that naturally binds the substrates or are very close to those sites. In fact, most of the compounds that are tested have affinity for the active site. Enzyme interface inhibitors usually affect the catalytic function as a consequence of the fact that the aggregation state often directly influences catalytic function itself as in the case of HIV protease [17]. Most of the work performed has not been able to show experimental details at the structural level to characterize the binding site of protein–protein interaction inhibitors [18]. The characterization of such interfaces represents the true barrier to effectively understanding the structural nature of PPIs and is consequently a major obstacle for designing modulating agents that can bind to those interfaces. In general, interaction interfaces are considered to be flat and rigid; in fact, they are dynamic and can be more convoluted in solution than they appear in co-crystal structures. This complexity causes some difficulties in the selection of the suitable template conformation. The properties that describe the interactions include the size and shapes of the interface, the accessible surface area, hydrophobicity, and the propensity of residues to participate in hydrophobic or hydrogen bond interactions.

In Fig. 2.1, an example of an interface binding pocket that usually binds other proteins and forms heterodimeric interfaces is shown [19]. The Nef protein of HIV was studied via a combined virtual and experimental screening method, and no previously known inhibitors could be used as a starting point in a structure-based research program that targets an SH3-binding surface of the HIV type I Nef protein. High-throughput docking and the application of a pharmacophoric filter, on the one hand, and the search for analogy, on the other hand, identified drug-like compounds that were further confirmed to bind Nef in the micromolar range.

Computational chemists have developed different tools designed to specifically study PPI inhibitors, including machine-learning-based models, which are mainly decision trees that use a dataset of known PPI inhibitors and of regular drugs to determine a global physicochemical profile for putative PPI inhibitors [20]. The

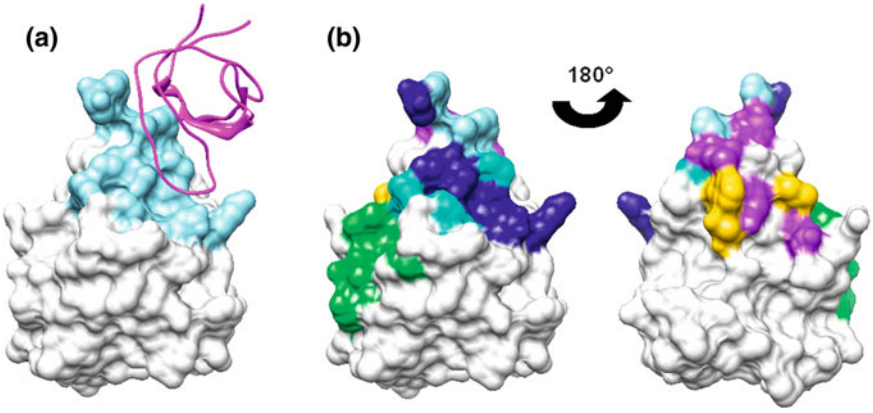


Fig. 2.1 HIV-1 Nef binding surfaces. **a** X-ray structure of the Nef–SH3Hck complex (Protein DataBank ID code: 1AVZ). The Nef surface is represented in light gray with the SH3 binding regio in light blue; the SH3 domain is shown in magenta ribbon. **b** Residues of the Nef core domain molecularly defined to be involved either directly or indirectly in cellular protein interactions and Nef functions are colored as follow: (i) SH3 binding region in blue ("proline-rich region" in sky blue, "RT loop binding region" in cyan, all the remaing residues in navy blue); (ii) p21 associated kinase 1/2 binding residues in purple; (iii) CD4 binding residues in green; (iv) Dyn2 and human thioesterase binding residues in yellow.

statistical analysis of the properties of effective PPI inhibitors included in the dataset, in relation to the compound's propensity to bind at the interface, unravels two important molecular descriptors for PPI inhibitors: specific molecular shapes and the presence of a privileged number of aromatic interactions that represent dominating aspects. These aspects are analyzed in another chapter within this book.

Attempts have been made to characterize the interacting interfaces in PPIs involved in inhibitors binding. A large variety of these PPI databases [21–25] depict PPIs at a structural level (for a summary of these available databases refer to [26]). A new system has been proposed by Morelli and collaborators [27] in which the specific PPI chemical space is represented through the presentation of the protein–protein interaction inhibitor (2P2I) database (DB), a hand-curated database dedicated to the structures of PPIs with known inhibitors (Fig. 2.2). PPIs and protein/inhibitor interfaces have been analyzed in terms of geometrical parameters, atom and residue properties, the buried accessible surface area, and other biophysical parameters. The interfaces found in 2P2I DB were then compared to those of representative datasets of heterodimeric complexes. In particular, they were compared with the interfaces found in representative datasets of heterodimeric complexes from Bahadur and Zacharias [28] or from the ProtorP parameters (http://www.bioinformatics.sussex.ac.uk/protorp/). A new classification of PPIs with known inhibitors into two classes is proposed, depending on the number of segments present at the interface and corresponding to either a single secondary structure element or a more globular interaction domain. 2P2I DB complexes share

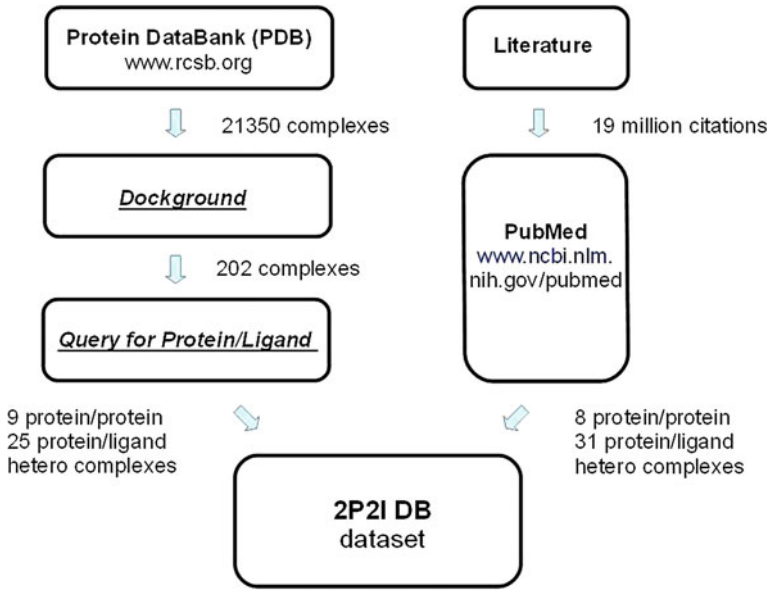


Fig. 2.2 2P2I database (*DB*), a hand-curated database dedicated to the structure of PPIs with known inhibitors (from [27])

global shape properties with standard transient heterodimer complexes, but their accessible surface areas are significantly smaller. No major conformational changes have been observed between the different states of the proteins. The interfaces are more hydrophobic than general PPI interfaces, with less charged residues and more non-polar atoms. Finally, 50 % of the complexes in the 2P2I DB dataset possess more hydrogen bonds than typical protein–protein complexes. Potential areas of study for the future have been proposed and include a new classification system consisting of specific families and the identification of PPI targets with high druggability potential based on key descriptors of the interaction.

However, proteins may have other roles, such as regulatory functions that modulate the expression of other proteins performed through the interaction of the protein with an mRNA molecule. This new generation of drugs can act as competitive antagonists or make much more subtle alterations through allosteric inhibitions, by only disrupting the way in which a protein interacts with other specific proteins (Fig. 2.1). These considerations suggest that the allosteric inhibitors or PPI interactions may generate new mechanisms of inhibition of the same protein with respect to active site inhibitors. In general, PPI inhibitor must have high affinity for monomer surface A and must repel monomer surface B where AB is the target dimeric protein. This concept is still not widely considered in the drug design approach of new PPIs inhibitors.

2.3 Structural Features of Protein–Protein Interaction Inhibitors

The emergence of PPI development programs has been driven by technological and conceptual advances in understanding PPI druggability and the types of molecules that can be used to block interactions. Biologics, in many ways, are more natural candidates for the inhibition of PPIs because they more likely resemble a natural PPI partner-like interface peptides mimic and their large size offers more opportunities to block an interaction and hit various hotspots (critical spots on the protein surface that are relevant for the interaction with a physiological partner protein). A large variety of compounds with nearly unrelated structures have been shown to interfere with the same protein–protein interface. These compounds may bind to the enzyme active site, if the active site is naturally located at the interface of the two or more monomers, they may bind to allosteric binding sites with respect to the substrate-binding sites. Their structures are not conserved, and there is no common general structure–activity relationship (SAR) that can recognize these compounds. However, systematic analysis has been performed to classify the protein–protein interaction inhibitors (2PP2Is). Morelli et al. [26, 27] reported that the 2P2I DB (http://2p2idb.cnrs-mrs.fr) contained a total of 17 protein/protein complexes corresponding to 14 families and 56 small-molecule inhibitors bound to the corresponding target (Fig. 2.3).

In the remaining part of this chapter, we will describe success stories within the PPI studies and natural compounds known as PPIs.

2.4 Successful and Recent Stories: γ-Secretase, Caspase 9, XIAP/BIR3, SMAC System, p21-Ras Oncoprotein, Human Thymidylate Synthase

Several recent success stories indicate that protein–protein interfaces might be more tractable than previously thought. These studies discovered small molecules that bind with drug-like potencies to “hotspots” on the contact surfaces involved in PPIs. Remarkably, these small molecules bind deeper within the contact surface of the target protein and bind with much higher efficiencies compared with the contact atoms of the natural protein partner. Some of these small molecules are now making their way through clinical trials.

2.5 γ-Secretase

γ-Secretase is a high molecular weight (HMW) membrane protein complex that includes presenilin (PS), nicastrin, Aph-1, and Pen-2. γ-Secretase is an

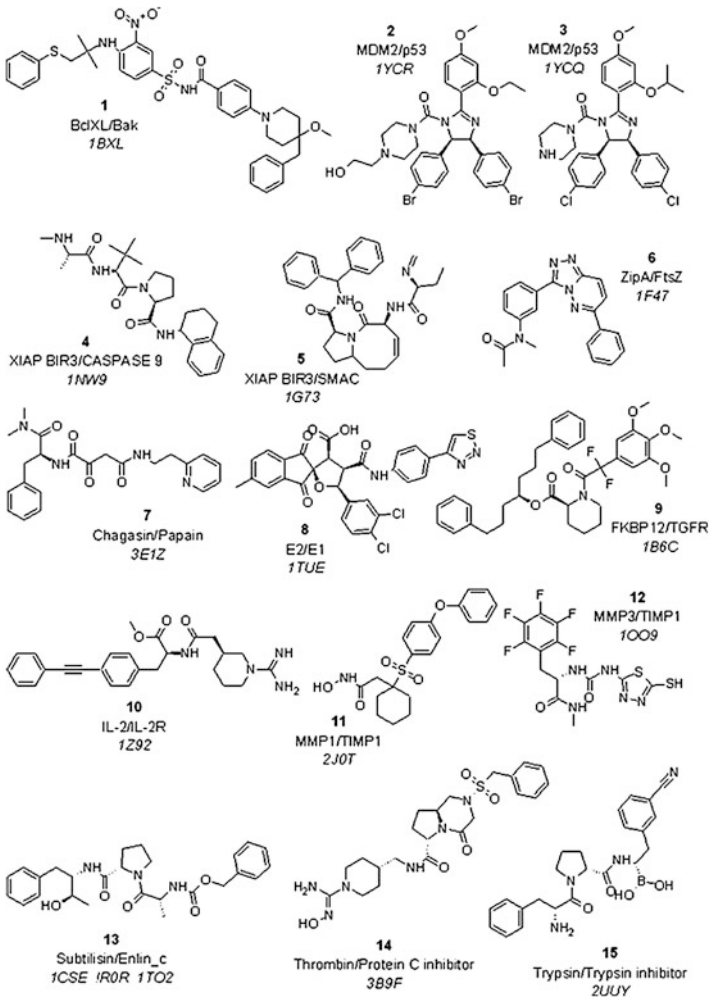


Fig. 2.3 List of representative small-molecule inhibitors for each protein–protein complex in the 2P2I DB. Only the inhibitor used to define the subset of the interface at 4.5 Å around the ligand is shown. For each inhibitor, the name of the protein family, the PDB code of the complex, and the molecular weight of the ligand are indicated

intramembrane-cleaving aspartyl protease [29]. The catalytic activity of the γ-secretase complex centers on the presenilin-1 (PS-1) protein, which is perhaps most notable for its role in mediating the cleavage of the β-amyloid (Aβ) peptide from the amyloid protein precursor (AβPP) [30]. AβPP is cleaved initially by α- or β-secretase to generate membrane-bound C-terminal fragments (C83 and C99, respectively). α-Secretase activity appears to be mediated by the disintegrin and metalloprotease (ADAM) family members TACE and ADAM-10. β-secretase (BACE, Asp-2) was recently cloned and characterized as a novel membrane-bound

aspartyl protease. The C83 and C99 cleavage products serve as substrates for the γ-secretase, and Aβ(40) and Aβ(42) are generated from C99. This cleavage is rather unusual in that it occurs at a site within the putative transmembrane domain of AβPP. Aβ42 is the major constituent of the amyloid plaques in the brain parenchyma of Alzheimer's disease patients (Fig. 2.4) [31].

The first class of γ-secretase inhibitors consisted of transition-state analogues designed to interact specifically with the diaspartyl active site of γ-secretase. Some of these inhibitors are peptidomimetics, which contain a non-hydrolyzable difluoro-alcohol moiety that resembles the transition state of the proteolysis carried out by aspartyl proteases or contain a non-hydrolyzable difluoro-ketone moiety that readily forms a hydrate resembling the aspartyl-protease-catalyzed transition state (Fig. 2.5). Moreover, it has been found that installing large, hydrophobic P1 substituents (such as *sec*-butyl or cyclohexylmethyl) into these transition-state mimics enhances their potency, indicating that there may be a relatively large complimentary S1 pocket in γ-secretase capable of accommodating these bulky substituents [32]. L-685,458 is a specific inhibitor of AβPP γ-secretase that contains a hydroxyethylene dipeptide isostere (Fig. 2.5) [31].

An α-helical peptide-based inhibitor has been described to bind PS fragments in a manner distinct from that of transition-state analogue inhibitors [29]. New inhibitor prototypes that would assume a helical conformation, such as short peptides based on the AβPP transmembrane domain, were designed. Evidence suggests that the AβPP transmembrane domain is helical upon its initial interaction with the γ-secretase complex. To achieve this conformational constraint, a well-known helix-inducing residue, α-aminoisobutyric acid (Aib), was incorporated. In this design, AβPP residues were judiciously swapped with Aib so that, upon helix formation, the Aib residues would be on one face of the helix and the APP residues would be on the other. Further, N-terminal *tert*-butoxycarbonyl, C-terminal methyl ester, and threonine *O*-benzyl protecting groups were retained to enhance cell permeability. These peptides blocked Aβ production from AβPP-transfected Chinese hamster ovary (CHO) cells in the low μM range, with longer peptides showing higher potencies. Inhibition occurred at the γ-secretase level, as AβPP γ-secretase substrates were increased in a concentration-dependent manner consistent with the effects on Aβ production. Surprisingly, the enantiomers of

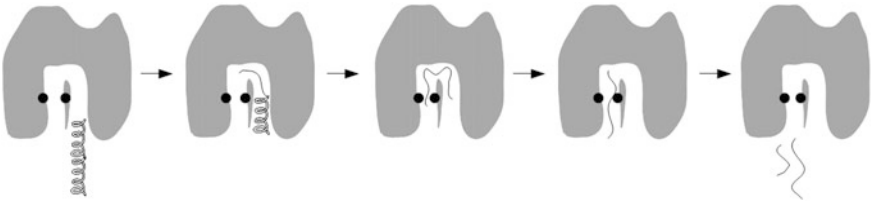


Fig. 2.4 Scheme of the enzymatic function of γ-secretase. The peptidic substrate carrying an α-helical transmembrane domain structure enters the "initial binding site"; then, it goes on to the catalytic site and undergoes an unwinding process. The substrates are then proteolyzed and the resulting fragments are liberated

these compounds, with all D-amino acids, were equally or more potent than their L-peptide counterparts because both enantiomers are helical in solution [33].

Finally, a number of structurally diverse and potent γ-secretase inhibitors have been reported in addition to the classical transition-state analogues. One of these non-transition-state analogue γ-secretase inhibitors is DAPT {N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester} (Fig. 2.5) [29]. This inhibitor was discovered as a result of a sandwich ELISA screen of approximately 25000 compounds for their ability to inhibit cellular Aβ production. A chemical optimization program was initiated around one active compound, N-arylalanine ester, which emerged from this screen [34]. Other compounds that were identified included arylsulfonamides and benzodiazepines [35].

It is important to mention that this HMW membrane protein complex seems to have different sites for catalytic function, substrate binding, and inhibitor binding. Recent studies based on inhibitor cross-competition kinetics indicated the presence on γ-secretase of two inhibitor binding sites, one for binding to transition-state isosteres and the other for non-transition-state small-molecule inhibitors [36]. Moreover, because γ-secretase cleaves amide bonds within the transmembrane regions of its substrates via a poorly understood process of hydrolysis within a hydrophobic environment and because of its requirement for water, the active site of γ-secretase is thought to be in the protein interior to avoid the hydrophobic environment of the lipid bilayer. As a result, the integral membrane substrates should initially interact on the surface of the protease before entering the internal active site. Substrate-binding sites that are distinct from the active site are generally called exosites or substrate-docking sites. It is believed that helical peptide inhibitors bind to γ-secretase at the substrate-docking site [37].

One compound, LY450139 or Semagacestat, a benzolactam γ-secretase inhibitor (Fig. 2.5), has entered clinical trials. These studies have shown a dose-dependent decrease in the plasma but not the cerebral spinal fluid levels of Aβ.

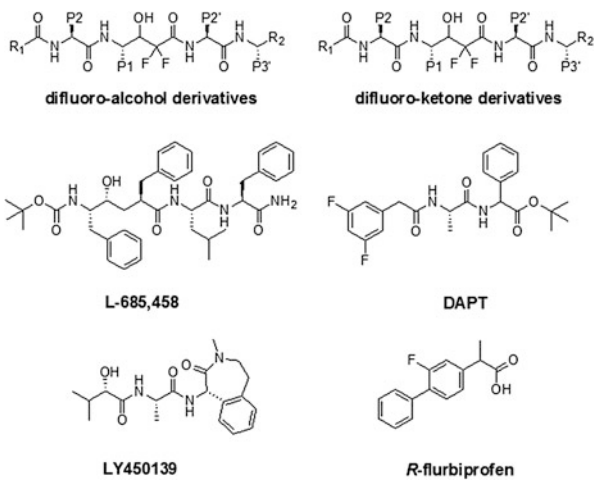


Fig. 2.5 Structures of select γ-secretase inhibitors

However, Semagacestat had no cognitive or functional benefit for patients with mild to moderate Alzheimer's disease in a multi-center, double-blinded, randomized, placebo-controlled phase II study. In addition, it is disappointing that preliminary results from two phase III clinical trials showed that Semagacestat did not slow down disease progression and was associated with a worsening of clinical measures of cognition and the ability to perform the activities of daily life. Unfortunately, Semagacestat is also associated with an increased risk for skin cancer. Consequently, the clinical trial of LY450139 was halted [38, 39].

Although γ-secretase is in many ways an attractive target for Alzheimer's disease therapeutics, interference with Notch processing and signaling may lead to toxicities that preclude the clinical use of inhibitors of this protease. Nevertheless, hope remains that a γ-secretase inhibitor might lower Aβ production in the brain enough to prevent Aβ oligomerization and fibril formation while leaving enough Notch signaling intact to avoid toxic effects. This hope has stimulated the continuing clinical studies of LY450139, even though compounds in this general structural class have not demonstrated selective inhibition of AβPP processing with respect to Notch [39–41].

In contrast, compounds that can modulate the enzyme to alter or block Aβ production with little or no effect on Notch would bypass this potential roadblock to the development of therapeutics. Recent studies suggest that the protease complex contains allosteric binding sites that can alter substrate selectivity and the sites of substrate proteolysis [40].

Certain non-steroidal anti-inflammatory drugs (NSAIDs) can reduce the production of the highly aggregation-prone Aβ42 peptide and increase the level of a 38-residue form of Aβ. This pharmacological property is independent of the inhibition of cyclooxygenase. Enzyme kinetics studies and displacement experiments suggest that selective NSAIDs can be non-competitive with respect to the AβPP substrate and to a transition-state analogue inhibitor, which also suggests an interaction with a site distinct from the active site and the docking site. While the biochemical mechanism of these NSAID γ-secretase modulators is unclear, the effect on AβPP proteolysis resulting in lower Aβ42 and increased Aβ38 has been well documented. The site of cleavage within the Notch transmembrane domain is similarly affected, but this subtle change does not inhibit the release of the intracellular domain and thus does not affect Notch signaling. For this reason, these agents may be safer as Alzheimer therapeutics than inhibitors that block the active site or the docking site. Indeed, one compound, *R*-flurbiprofen (formerly flurizan, now called tarenflurbil) (Fig. 2.5), has recently advanced to phase III clinical trials in the USA. The enantiomer, *S*-flurbiprofen, is a COX inhibitor, whereas *R*-flurbiprofen modulates Aβ production but is not a COX inhibitor. However, the potency of this drug candidate and other NSAIDs in terms of lowering Aβ42 raises questions about efficacy, and it will likely be important to better understand how these compounds interact with PS and develop second- and third-generation agents [40].

Another type of allosteric modulator includes compounds that resemble kinase inhibitors and interact with a nucleotide-binding site on the γ-secretase complex.

The discovery that ATP can increase Aβ production in membrane preparations prompted the testing of a variety of compounds known to interact with ATP-binding sites on other proteins. In a focused screen, the kinase inhibitor Gleevec emerged as a selective cellular inhibitor of Aβ production without affecting the proteolysis of Notch. In light of these findings, ATP and other nucleotides were tested for the effects on purified γ-secretase preparations and were found to selectively increase the proteolytic processing of a purified recombinant AβPP-based substrate without affecting the proteolysis of the Notch counterpart. This and other evidence suggest that the γ-secretase complex contains a nucleotide-binding site and that this site allows the allosteric regulation of the γ-secretase-mediated processing of AβPP with respect to Notch. Whether this regulation is physiologically important is unclear, but the pharmacological relevance is profound and may lead to the development of new therapeutic candidates for AD. This hope is tempered by the fact that γ-secretase cleaves numerous other type I membrane protein substrates. The use of agents that are selective for APP *versus* Notch may result in new long-term toxicities as a result of blocking the proteolysis of other substrates. To address this important issue, the development of more potent analogues that work by this mechanism will be critical [40].

2.6 Caspase 9, XIAP/BIR3, SMAC System

Apoptosis, or programmed cell death, originates from at least two different pathways: an intrinsic pathway initiated from mitochondrial membrane depolarization as a result of irreparable damage to the genome, microtubule disruption, and other factors, and an extrinsic pathway activated by tumor necrosis factor family death receptors. Both apoptotic pathways culminate in the activation of cysteine protease family members known as caspases. Caspase-9 is an initiator caspase in the intrinsic pathway. Upon the recruitment to a multiprotein structure called the apoptosome, procaspase-9 dimerizes into a catalytically active form that cleaves and activates procaspase-3 and procaspase-7 (Fig. 2.6) [42].

Apoptosis plays a crucial role in the homeostasis and development of living organisms. Deregulation of this mechanism is associated with many diseases, including several types of cancer.

In the apoptosis pathway, the inhibitors of apoptosis proteins (IAPs) are exploited by tumor cells to evade programmed cell death. The XIAP (X-linked inhibitor of apoptosis protein) is the most potent caspase inhibitor among the IAP protein family. XIAP contains three baculoviral inhibitory repeat (BIR) domains and a ring domain. This protein interacts with initiator caspase-9 and executioner caspase-3 and caspase-7 through its BIR3 and BIR2 domains, respectively. BIR3 inhibits caspase-9 by preventing dimerization, which is required for its catalytic activity. The search for new compounds that are able to disrupt the XIAP–caspase interaction has attracted the attention of scientific community as a promising strategy for cancer treatment (Fig. 2.6) [42, 43].

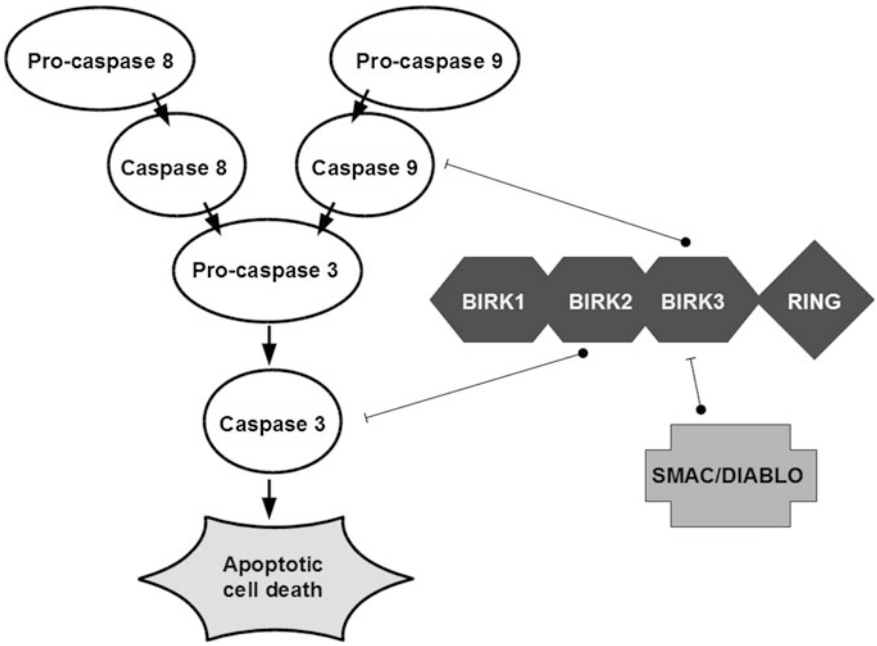


Fig. 2.6 Two different pathways to the programmed cell death, that is, apoptosis

Two broad approaches have been taken to develop clinical inhibitors of XIAP—antisense oligonucleotides and small-molecule inhibitors. Antisense oligonucleotides, aside from the disadvantages inherent to the method, are advantageous, as they target the entire protein, whereas small molecules bind a single domain. However, small-molecule inhibitors offer the potential of a more rapid inhibition of their target in vivo and a more predictable duration of action.

Antisense oligonucleotides directed against XIAP are being developed for the treatment for solid tumors and hematologic malignancies by Aegera pharmaceuticals (Montreal, Canada). The antisense molecule currently in clinical trials (AEG 35156) is a second-generation 19-mer antisense oligonucleotide that targets XIAP. It contains a mixed backbone of chemically modified DNA/RNA nucleotides. Antisense oligonucleotides inhibit their target by forming duplexes with the intracellular native mRNA. The duplexes recruit RNase H enzymes that cleave the native mRNA strand while leaving the antisense oligonucleotide intact. The antisense oligonucleotide is then released back into the cytosol where it is capable of inhibiting additional native mRNA molecules. In cultured cells, the uptake of antisense oligonucleotides requires transfection, infection, or electroporation protocols. In xenografts and patients, the intracellular uptake of antisense molecules is achieved after intravenous or subcutaneous administration, but the mechanism facilitating the uptake is not clear.

Although the xenograft data have been encouraging, the efficacy of antisense therapies in previous clinical trials has not matched the initial expectations. To

date, most antisense trials have indicated that these antisense oligonucleotides are safe, but their efficacy is modest. Given the redundancy in the IAP family, the selective targeting of XIAP may not produce an optimal effect. In this case, a different approach using pan-IAP inhibitors may be needed. The lack of a significant response to antisense oligonucleotides may also reflect a failure to achieve adequate target knockdown, as it is unknown how much knockdown is required to inhibit the function of the target. Finally, it remains unclear whether the effects of antisense therapies are related to the knockdown of their target or non-specific effects resulting from altered gene regulation. Given these concerns, XIAP antisense therapy may also be hampered in clinical trials. Therefore, there is great interest in developing small-molecule inhibitors of XIAP to overcome the limitations of antisense oligonucleotides [44].

The natural inhibitor of XIAP, SMAC/DIABLO (second mitochondrial activator of caspases/direct IAP-binding protein with Low PI), is a protein that is released from the mitochondria into the cytosol in response to apoptotic stimuli. SMAC removes the XIAP inhibition of caspase-9 by binding to the BIR3 domain of XIAP through the AVPI tetrapeptide present in the N-terminal part of SMAC. This interaction (AVPI/BIR3) has been determined unequivocally by X-ray crystallography [43].

Since the discovery of the SMAC protein in 2000, there has been an enormous interest in academic laboratories and pharmaceutical companies in the design of small-molecule SMAC mimetics [45]. Various design methods have been used to generate peptidomimetics and non-peptide SMAC mimetics with improved stability and enhanced drug-like properties. A number of research groups have reported the discovery of small-molecule BIR3 inhibitors by various methods, including peptidomimetic approaches, virtual screening/structure-based design, or the screening of natural products or synthetic libraries [46].

Using the AVPI structure (Fig. 2.7), a comprehensive study has been performed to determine which amino acids could be substituted without compromising the peptide-binding affinity. The alanine (first amino acid) and the proline (third amino acid) are the essential amino acid residues needed to preserve the activity of this tetrapeptide. However, the use of synthetic SMAC-derived peptides as therapeutic compounds is hindered by their limited cell permeability, proteolytic instability, and poor pharmacokinetics [43]. To address these limitations, a number of early studies employed a strategy to tether a carrier peptide to a SMAC-based peptide to facilitate intracellular delivery. It was shown that these relatively cell-permeable SMAC-based peptides can sensitize various tumor cells in vitro to the anti-tumor activity of Apo-2L/TRAIL, as well as chemotherapeutic agents such as paclitaxel, etoposide, SN-38, doxorubicin, and cisplatin [45].

Extensive chemical modifications of the AVPI peptide have been carried out in an effort to derive potent SMAC peptidomimetics. Furthermore, conformationally constrained, bicyclic SMAC mimetics were designed using a structure-based strategy (Fig. 2.7) [45].

It has been demonstrated that the natural SMAC protein forms a dimer and binds to XIAP protein constructs containing BIR2 and BIR3 domains with a much

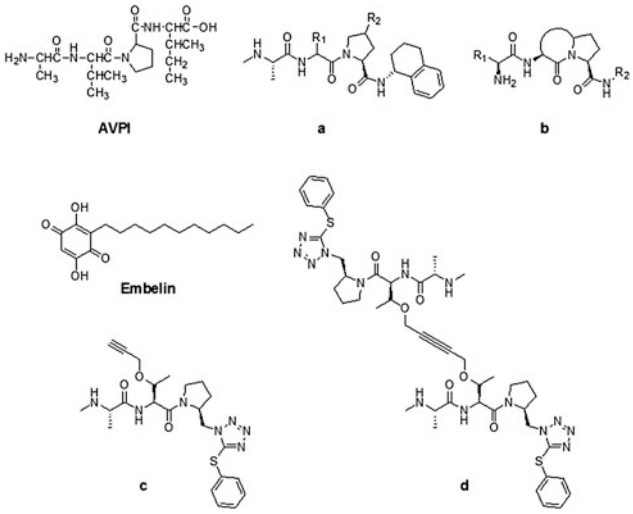


Fig. 2.7 Structure of some XIAP/BIRK3 inhibitors. AVPI peptide; some SMAC peptidomimetics (**a**); some conformationally constrained non-peptidic SMAC mimetics (**b**); Embelin, a monovalent SMAC mimetic (**c**); a bivalent SMAC mimetic (**d**)

higher affinity than the SMAC AVPI peptide. Indeed, functional studies have shown that SMAC protein is a much more efficient and potent antagonist than the AVPI peptide against XIAP protein containing the BIR2 and BIR3 domains in terms of relieving the XIAP-mediated inhibition of caspase-9, caspase-3, and caspase-7 activity. The SMAC AVPI-binding motif binds to both the BIR2 and BIR3 domains, although with a stronger affinity to BIR3. Thus, small molecules designed to have two "AVPI"-binding motifs may mimic the mode of action of SMAC protein to target XIAP and be capable of achieving very high binding affinities to XIAP by concurrently targeting both the BIR2 and BIR3 domains in the protein. The structure-based design of non-peptidic, bivalent SMAC mimetics based upon conformationally constrained monovalent SMAC mimetics has been reported. The study characterized in detail the interaction of both monovalent and bivalent SMAC mimetics with different XIAP protein constructs. As potential drug candidates, there are advantages and disadvantages associated with monovalent and bivalent SMAC mimetics (Fig. 2.7). Monovalent SMAC mimetics are less potent than their corresponding bivalent SMAC mimetics. However, monovalent SMAC mimetics, with a molecular weight of $\sim$500 Da, possess many desirable pharmacological properties as potential drug candidates. Bivalent SMAC mimetics have been shown to be 100–1,000 times more potent than their monovalent counterparts and could potentially be far more efficacious. However, because bivalent SMAC mimetics have a molecular weight exceeding 1,000 Da, such compounds may be expected to have very low oral bioavailability and will have to be administered by other routes of administration, such as intravenous dosing,

which represents a potential disadvantage if the drug must be given to patients frequently [45].

In 2004, embelin, a natural compound(Fig. 2.7) from the Japanese Ardisia herb, was reported as a fairly potent, non-peptidic, small molecular weight, cell-permeable inhibitor that binds to the XIAP BIR3 domain. Embelin was identified through structure-based computer screening of the proprietary searchable three-dimensional structure database (TCM-3D) containing 8,221 small organic molecules with diverse chemical structures isolated from nearly 1,000 traditional Chinese medicinal herbs. Unlike most commercial databases, all of the compounds in the TCM-3D database are natural products derived from traditional medicinal herbs, which have been used for medicinal purposes in China and other countries for centuries. The extensive use of these traditional Chinese medicine recipes in humans has generated a great amount of data about their efficacy and safety [47].

2.7 p21-Ras

Ras proteins play a central role in cellular growth and differentiation. Ras signaling is tightly regulated by cycling between an active GTP-bound conformation (Ras-GTP) and an inactive GDP-bound state (Ras-GDP) (Fig. 2.8).

Ras has a slow intrinsic GTPase activity that is enhanced by GTPase-activating proteins (GAPs). These proteins greatly increase the rate of GTP hydrolysis and thereby act as negative regulators of Ras output. Neurofibromin also regulates Ras signaling pathways by accelerating the conversion of Ras-GTP to Ras-GDP. Oncogenic *RAS* alleles carry single-point mutations at the Gly12, Gly13, or Gln61 positions that greatly reduce its intrinsic GTPase activity and render these proteins resistant to GAPs. Oncogenic *RAS* mutations or the inactivation of *NF1* perturb Ras signaling by favoring the GTP-bound state (Fig. 2.8).

Growth factors induce cell growth, in part by activating guanine nucleotide–exchange factors (GEFs), such as Sos, GRF, and GRP, which bind Ras and stimulate nucleotide dissociation. Nucleotide exchange increases the percentage of Ras-GTP because the intracellular concentration of free GTP vastly exceeds that of GDP. Signaling terminates when Ras-GTP is hydrolyzed to Ras-GDP (Fig. 2.8).

Ras proteins regulate cell fates by transducing signals from the plasma membrane to the nucleus via a series of downstream effectors. Ras-GTP recruits Raf kinase to the membrane, where it initiates a kinase cascade involving MEK kinase and the Erk1 and Erk2 isoforms of mitogen-activated protein (MAP) kinase. The activation states of the phosphoinositol-3′-kinase (PI3K) and Rac/Rho pathways are also regulated by Ras-GTP in many cell types. The consequences of Ras activation are influenced by the cellular context and by cross-talk between its downstream effectors. Ras-GTP-mediated activation of the Raf/MEK/ERK kinase cascade stimulates proliferation in many cell types, and activation of the PI3K pathway has been shown to promote cellular survival. Ras also interacts with other downstream effectors, such as Ral-GDS.

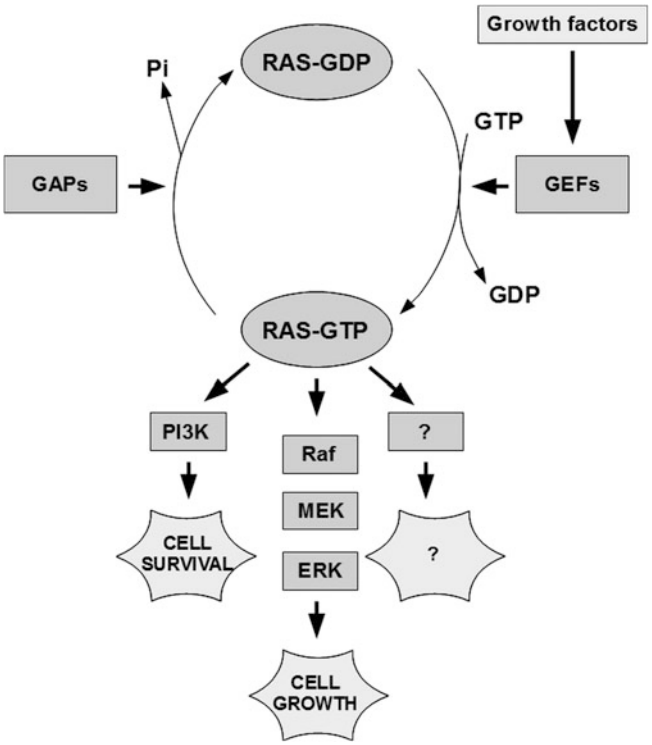


Fig. 2.8 Schematic representation of Ras cycling between the active *GTP-bound* conformation and the inactive *GDP-bound* state, the cascade effects of Ras activation and modulation by *GEF*s and *GAP*s

The posttranslational modification of Ras is initiated by the attachment of either a farnesyl or geranylgeranyl isoprenoid lipid to the cysteine residue of the carboxy-terminal CAAX box. These reactions are catalyzed by farnesyl and geranylgeranyltransferase, respectively. The last three amino acids (i.e., the -A-A-X) are then removed by a specific endoprotease, Rce1. The final step in Ras processing involves methylation of the carboxyl group by an endoplasmic reticulum–associated methyltransferase [48].

Recent studies suggest that the p21 product of the Ras oncogene may be an obligatory intermediate in transducing the growth factor signal. The activation of Ras may, therefore, activate the growth factor pathway without the need for either a growth factor or its receptor [49].

Functionally, all p21-Ras isotypes can be divided into three domains: (1) the NH2-terminal portion, which contains the phosphate-binding and effector domains and has almost 100 % homology between isotypes; (2) the intermediate guanine recognition domain, which also has high homology; and (3) the COOH-terminal hypervariable region, which undergoes posttranslational modification, including farnesylation, leading to membrane localization. The crystal structure of inactive

and activated p21-Ras has also revealed two "switch" domains, in which the tertiary structure of the protein alternates between inactive and active states. The switch-1 domain corresponds to the effector domain in the NH2 terminus. The conformation of this domain is altered with activation, allowing the interaction of downstream effectors. The switch-11 domain is similar to an intermediate guanine nucleotide recognition domain, which assumes a different conformation that is dependent on the nucleotide exchange between GDP and GTP caused by the action of Ras-GEFs and Ras-GAPs. Data from recent mouse knockout studies have demonstrated that, despite their high degree of homology, p21-Ras isotypes do possess certain unique cellular functions [50].

Because efforts to identify and develop small-molecule inhibitors that directly target Ras have not been successful, a majority of past and ongoing efforts have targeted Ras indirectly, to modulate the functions of proteins that influence or mediate Ras oncogenesis.

The low micromolar-binding affinity of protein kinases for ATP, where potent nanomolar affinity ATP-competitive inhibitors have been developed (e.g., imatinib), has been a very successful avenue for anti-cancer drug development. In contrast, the low picomolar-binding affinity of small GTPases for GTP and millimolar cellular concentrations of GTP renders a similar strategy for Ras implausible. Thus, past and current efforts have focused on indirect approaches for the disruption of Ras function, that is, the inhibition of components that regulate Ras membrane association and the inhibition of downstream effector signaling.

One of these classes of proteins includes those that regulate Ras posttranslational processing, which is either signaled through the C-terminal CAAX tetrapeptide motif (farnesyltransferase, FTase; Rac-converting enzyme 1, Rce1; Isoprenylcysteine carboxyl methyltransferase, Icmt) or by protein kinase C alpha (PKCα)-dependent phosphorylation or ubiquitination.

Considerable past efforts centered on the development of FTase inhibitors (FTIs), with many identified and two remaining in clinical trial analyses (lonafarnib and tipifarnib, Fig. 2.9). The prenylation of K-Ras and N-Ras by a related enzyme, geranylgeranyltransferase-I, when farnesyltransferase activity is blocked by treatment with an FTI, proved to be the downfall of FTIs as effective Ras inhibitors [51, 52].

A second class of inhibitor of Ras membrane association is composed of two small molecules with farnesyl lipid groups (salirasib and TLN-4601, Fig. 2.9) that are proposed to compete with Ras for membrane-associated docking proteins for the Ras isoprenoid group. Efforts to target Ras effector signaling first centered on the Raf–MEK–ERK MAPK cascade. Small-molecule protein kinase inhibitors of MEK1/2, and later Raf, have been developed, with many inhibitors now under clinical evaluation. More recently, inhibitors of the p110 catalytic subunits of PI3K, AKT, and mTOR have entered clinical trials, and two mTOR inhibitors have been approved by the FDA for the treatment for renal cell cancers [52]. Moreover, Ras proteins are also regulated by multiple GEFs and GAPs. GTP-bound Ras interacts with catalytically diverse downstream effectors that possess

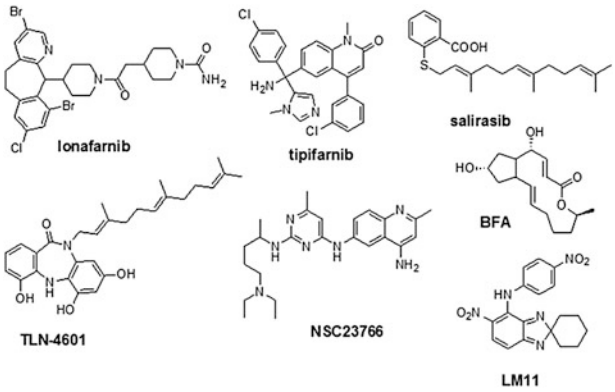


Fig. 2.9 Structures of some FTase inhibitors (lonafarnib, tipifarnib), inhibitors of Ras membrane association (salirasib, TLN-4601), and GEF inhibitors (BFA, LM11, NSC23766)

Ras-binding (RBD) or Ras association (RA) domains. Some of these interactions are restricted to specific Ras isoforms [52].

Brefeldin A (BFA, Fig. 2.9) is a natural product isolated from the fungus *Eupenicillium brefeldianum* and is the first known inhibitor of a GEF. The extensive analysis of the mechanism of action of BFA led to the general concept of "interfacial inhibition," which refers to inhibitors that act by stabilizing protein complexes and target regions in or near interfaces. Some inhibitors of natural origin that are already used in the clinic have been recognized as interfacial inhibitors, such as the anti-cancer drugs vinblastine or camptothecin. LM11 (Fig. 2.9) was discovered by an in silico screen and was shown to target an interfacial depression at the surface of the complex between Arf1–GDP and BFA-insensitive GEFs, such as ARNO, and to block ARNO-dependent cellular migration. A few other promising examples of cell-active small-molecule ArfGEF inhibitors have been selected by in vitro and phenotypic screens. These studies demonstrate that despite the high levels of homology that are found within a given GEF family, specific GEF inhibitors can be developed. Therefore, the specific flexibility and conformational changes that characterize small GTPase–GEF complexes are likely to be advantageous to drug development, notably for interfacial inhibitors. However, the design of high-throughput biochemical assays to effectively screen for such inhibitors remains a challenge. Inhibitors that target specific RhoGEFs have been discovered by high-throughput screens.

Another related example of a way to target GTPase activity is through targeting the surface of GTPases that is required for GEF activation. The small-molecule NSC23766 (Fig. 2.9) was discovered through computational screening of the surface of Rac1, which is known to interact with GEFs. This compound was found to inhibit the activation of Rac1 by the Rac-specific GEFs, Trio, and Tiam1.

There is limited but promising evidence that it may be possible to develop small-molecule modulators of Ras superfamily GAPs. HTS identified small-molecule inhibitors of RGS domains, which are GAPs for heterotrimeric G proteins.

Despite their low structural homology to RasGAPs, they share a similar enzymatic transition state, suggesting that this could be a starting point for the design of Ras superfamily GAP inhibitors. The Rac-selective chimaerins (CHN), one class of RhoGAPs, possesses C1 zinc finger domains that bind diacylglycerol, a cofactor for their activity. Therefore, small molecules that bind C1 domains may activate their GAP activities, causing the downregulation of Rac GTPase activity. While such a therapeutic approach will be complicated by the existence of other proteins with C1 domains (e.g., RasGRP), there is evidence that C1-binding molecules can have some degree of selectivity for a subset of C1-containing proteins. This approach may be a therapeutic option for cancers in which there is RacGEF-mediated activation of Rac [52].

Another approach to antagonize Ras functions in vivo is the inhibition of Ras expression by anti-sense RNA-based technology, the use of anti-Ras ribozymes (hammerhead models), or the application of anti-Ras retroviral therapy [53].

2.8 Human Thymidylate Synthase

Human thymidylate synthase (hTS) is an enzyme that plays a key role in DNA synthesis and is a target for several clinically important anti-cancer drugs. Inhibitors of hTS are widely used in chemotherapy; the best known inhibitors are raltitrexed, pemetrexed, and 5-fluorouracil [54]. However, their use is associated with drug resistance. Therefore, compounds with different inhibitory mechanisms are required to combat resistance. Some peptides have been designed to specifically target the protein–protein interface in hTS, which is a dimeric protein composed of two identical polypeptide chains. The peptides stabilize the inactive form of the enzyme and inhibit cell growth in drug-sensitive and drug-resistant cancer cell lines.

TS (EC: 2.1.1.45) catalyzes the reductive methylation of 2′-deoxyuridine-5′-monophosphate (dUMP) to 2′-deoxythymidine-5′-monophosphate (dTMP), which is assisted by the cofactor N^5, N^{10}-methylene tetrahydrofolate (MTHF) [55–57]. Because TS represents the only synthetic source of dTMP in cells, it is a major target for the design of chemotherapeutic agents [58].

In addition to its catalytic role, hTS has also been shown to regulate protein synthesis by interacting with its own messenger RNA as well as those of several other proteins involved in the cell cycle [59]. The regulatory function of hTS as an RNA-binding protein has been shown to be maximal when the protein is not bound by ligands. This observation, together with the observation that cancer cells resistant to anti-hTS drugs, has increased levels of hTS, led to the hypothesis that the overexpression of hTS might be correlated with the loss of RNA regulatory capacity when the protein is bound to its inhibitors [60]. In the case of ovarian cancer, for example, access to therapy has been limited by the cross-resistance of platinum drugs with classical drugs that target the metabolic pathway that uses folic acid and its derivatives as substrates and/or cofactors. The same observation

may be valid for other cancer types, such as colorectal cancer, in which TS inhibitors are widely used. Although the overexpression of hTS has been observed in platinum-sensitive cells, this effect is even more pronounced in platinum-resistant ones. Therefore, it is important to identify hTS inhibitors that act through new mechanisms that do not alter RNA regulation or increase protein levels [61]. When Schiffer and coworkers [62] first crystallized the unliganded form of hTS, they found that the active site loop (residues 181–197) containing the catalytic cysteine (Cys195) was twisted approximately 180 degrees compared to the corresponding loop conformation in the liganded hTS. Because in unliganded hTS, Cys195 is outside the active site, the enzyme must be inactive (Fig. 2.10). The authors suggested that the inactive conformation might serve to protect the catalytic cysteine from cellular modification. Three phosphate/sulfate ions were observed to be bound near the active site, suggesting that inactive hTS can bind to TS mRNA, thereby repressing TS protein synthesis. In addition, the disordered small domain (residues 107–128) of the inactive hTS likely increases its proneness to cellular degradation, further reducing cellular TS levels. They demonstrated through fluorescence studies that there is an equilibrium between the active and inactive states; phosphate ions were shown to shift the equilibrium toward the inactive state, and the binding of dUMP causes a shift toward the active state [63]. The R163K mutant, which stabilizes the active conformer, is at least 33 % more active than wild-type hTS, suggesting that at least 1/3 of hTS exists in the inactive state [64]. On the basis of these data, some diphosphonic acids have been proposed as inhibitors of hTS. The most active one is 1, 3-propanediphosphonic acid (PDPA), which binds at the position where the phosphate group of the substrate dUMP binds. However, the mechanism of inhibition of these compounds has not yet been biochemically and mechanistically demonstrated. Furthermore, PDPA has no cellular activity, and thus its potential as a drug is limited.

On the basis of structural considerations, PDPA was discovered as an allosteric inhibitor of hTS. It has been proposed that PDPA acts by stabilizing an inactive conformer of loop 181–197. Kinetic studies showed that PDPA is a mixed (non-competitive) inhibitor versus dUMP. In contrast, PDPA is an uncompetitive inhibitor versus MTHF at concentrations lower than 0.25 μM, but at PDPA concentrations higher than 1 μM, the inhibition is non-competitive, as expected. At the concentrations corresponding to uncompetitive inhibition, PDPA shows positive cooperativity with an antifolate inhibitor, ZD9331, which binds to the active conformer. PDPA binding leads to the formation of hTS tetramers. These data are consistent with a model in which hTS preferentially exists as an asymmetric dimer with one subunit in the active conformation of loop 181–197 and the other in the inactive conformation.

A new area of investigation is the discovery of several peptides [65], with sequences from the protein–protein interface of the hTS dimer that inhibit hTS by a mechanism that involves selective binding to a novel binding site at the dimer interface of the inactive form of the enzyme. The peptides were designed to contain a sequence that is complementary to the monomer surface, and therefore, these peptide inhibitors are protein–protein interaction modulators. The

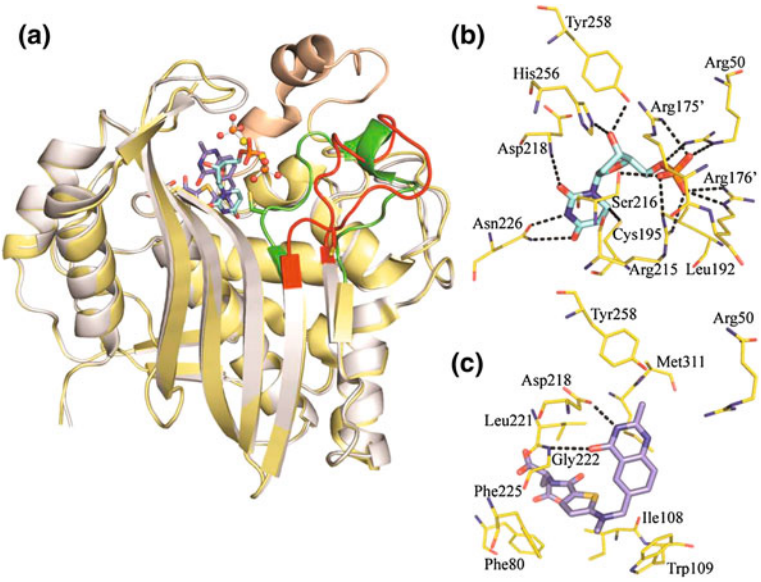


Fig. 2.10 **a** Cartoon representation of the superimposed monomeric subunits (from the crystallized dimers) of human thymidylate synthase in active (PDB 1HVY) (*yellow*) and inactive (PDB 2ONB) (*gray*) conformations. The active conformation of the active site loop is shown in green and the inactive conformation in *red*. The catalytic cysteine, C195, is highlighted with a stick representation on the loops. The small domain visible in the active crystal structure is shown in *brown*. dUMP (cyan sticks) and a folate analog, raltitrexed (*dark blue sticks*), are present in the active site, whereas PDPA (ball-and-stick representation) is located in an allosteric position. **b** Interactions of dUMP with the protein. *Dashed lines* represent direct hydrogen bonds between amino acid residues and the ligand, the *solid line* between Cys195 and dUMP represents a covalent bond. **c** Interactions of the folate analog, raltitrexed, with the protein

mechanism of inhibition has been demonstrated through a combination of experimental and computational approaches. Kinetic analyses, isothermal titration calorimetry, fluorescence spectroscopy, X-ray diffraction, and modeling studies show that this mechanism differs from those of the protein–protein interface inhibitors reported to date [2] because it involves the stabilization of an inactive form of the catalytic protein. The kinetic analysis suggests that these peptides behave as mixed-type inhibitors. The X-ray crystal structure of the hTS–LR peptide complex unambiguously shows that LR binds at the monomer–monomer interface and that the two monomers of the enzyme must move apart for the peptide to bind properly. However, Foerster resonance energy transfer (FRET) experiments show that the hTS monomers do not dissociate upon peptide binding. Calorimetric results indicated that the peptides only bind to one of the isoforms of the enzyme and that no binding is observed when the protein is saturated with the substrate dUMP, which is known to convert the protein to its active form. On the basis of these data, an inhibition model characterized by the stabilization of the inactive form of the enzyme by peptide binding has been proposed. This model is consistent with physical and biochemical observations.

Unlike the existing drugs that target hTS, these peptides inhibit intracellular hTS as well as cell growth without resulting in overexpression of hTS when administered to ovarian cancer cells. The connection between the stabilization of the inactive form of hTS by the peptides and the cellular effects remains to be fully explored. Further steps will require the optimization of the compounds via the synthesis of peptidomimetics and a detailed analysis of their cellular mechanism of action. The study demonstrates a mechanism in which a multifunctional homodimeric protein is inhibited through the binding of a ligand to the dimer interface of the inactive form of the enzyme, resulting in its stabilization. The concepts revealed here can be exploited to provide new strategies for the development of drugs for combating diseases such as ovarian cancer.

2.9 Natural Compounds as PPI Inhibitors

Several screening studies for the evaluation of the capacity of natural products to inhibit PPIs have been described in the literature. Natural product chemistry offers several advantages in drug discovery, mainly due to the great variety of molecular sizes and chemical structures provided by secondary metabolites isolated from plant species, fungi, microorganisms, marine invertebrates, algal species, and sponges. In particular, targeting PPIs with small natural compounds has been demonstrated to be a valuable application in the scope of anticancer drug discovery, as shown by several screening studies reported in the literature.

A first example is represented by a screen of approximately 7,000 purified natural compounds at 10 μM using high-throughput ELISA to identify molecules that inhibit the association between T-cell factor 4 (Tcf4) and *β*-catenin, which represents a rational and feasible target for the development of drugs against colorectal cancer [66]. Eight compounds resulted in a reproducible dose-dependent inhibition of the Tcf4/*β*–catenin interaction, with $IC_{50} < 10$ μM [66]. The group of active compounds had varying chemical structures, although some of them share polyhydroxylated planar features. Three compounds were isolated from fungal organisms, whereas three originated in actinomycete strains. One of the active molecules was isolated from a marine sponge, whereas the origins of the remainder were not specified. Further assays carried out to demonstrate that the isolated compounds work as predicted indicated that two fungal derivatives, named PKF115-584 and CGP049090 (Fig. 2.11), were able to disrupt the Tcf4/*β*–catenin complex in vitro and to inhibit colon cancer proliferation [66]. These antagonists of the Tcf4/*β*-catenin complex were also able inhibit the proliferation of adrenocortical carcinoma [67] and hepatocellular carcinoma [68] cell lines.

Another example is the study of Tsukamoto et al. [69], who described the inhibition of the p53/HDM2 interaction using a simple unsaturated dicarboxylic acid named (–)-hexylitaconic acid (Fig. 2.11), which was isolated from a culture broth of the marine-derived fungus *Arthrinium* sp. HDM2 (human double minute 2) or MDM2 (mouse double minute 2) is a ubiquitin ligase that induces the

degradation of p53, a tumor suppressor protein [70]. Targeting HDM2 or MDM2 is a useful tool to activate p53 and induce growth arrest and apoptosis in cancer cells [70]. The isolated natural compound was able to inhibit the p53/HDM2 interaction, and the IC_{50} value was found to be 50 μg/mL [69], as determined by ELISA. A more functionalized and complex natural product with the same activity is chlorofusin [71]. Chlorofusin is a highly modified cyclic peptide isolated from a microfungus of the genus *Fusarium*, which was demonstrated to inhibit the p53/MDM2 interaction with an IC_{50} value of 4.6 μM [71].

A further example of the capacity of natural compounds to inhibit PPIs is provided by Hashimoto et al. [72]. The authors developed a HTS system using an in vitro protein fragment complementation assay (PCA) to test the possible inhibition of the interaction between T-cell factor 7 (TCF7) and *β*-catenin, proteasome assembling chaperone (PAC)1 and PAC2, and the self-association of PAC3 homodimer, each of which play an important role in the growth of cancer cells, from 123,599 samples in a natural product library. A fungal metabolite named

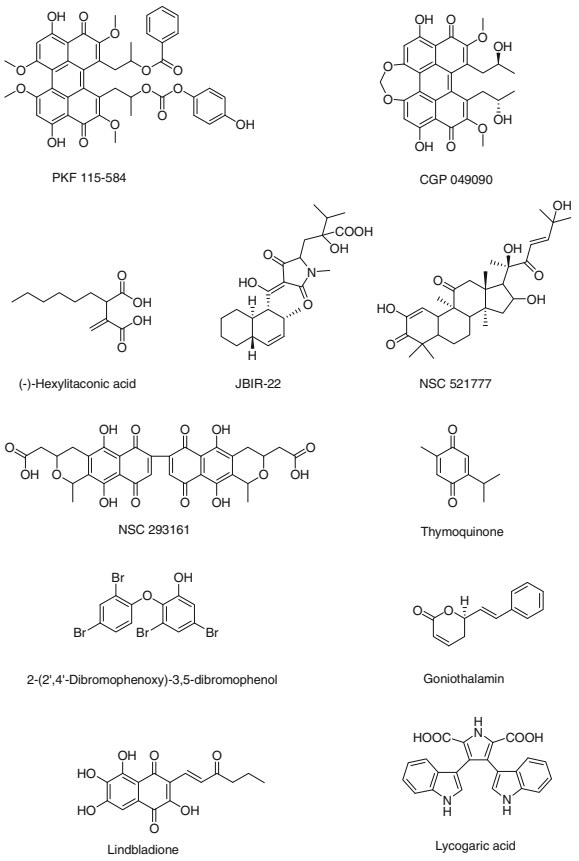


Fig. 2.11 Chemical structures of natural compounds able to inhibit PPIs

TB1, whose structure was not shown, was found capable of strongly and specifically inhibiting PAC3 homodimerization, with an IC_{50} value of 0.020 μM and was recognized as an invaluable lead compound for the development of new anticancer drugs. Another compound named JBIR-22 (Fig. 2.11) showed a specific inhibitory activity on PAC3 homodimerization, with an IC_{50} value of 0.2 μM [72]. This compound is a fungal metabolite isolated from *Verticillus* sp., whose structure was correctly clarified in a later study [73] and was revealed to be a new equisetin analogue on the basis of extensive NMR and MS analyses. Docking studies of this secondary metabolite with the X-ray structure of PAC3 revealed that it can bind to the active site of the PPI required for PAC3 homodimerization. In the same study [73], the cytotoxic activity of the isolated compound in the human cervical carcinoma cancer cell line (HeLa) was evaluated. The tested secondary metabolite displayed an IC_{50} value of 68 μM at 120 h, showing no effect at the same concentration after 48 h [73]. These results suggest that it may inhibit the generation of proteasome components.

In another study, a virtual screening of molecules present in the National cancer institute (NCI) molecular database, known as "NCI-diversity," which features approximately 2,000 anticancer agents based on fairly different chemical scaffolds, identified ten compounds, two of which were of natural origin, as inhibitors of stem cell factor (SCF) [74]. SCF is an endogenous growth factor involved in the hematopoietic cell proliferation and differentiation [74]. It also plays a crucial role in the development of melanoma and several intestinal cancers [74]. Similar to other growth factors, SCF dimerization is a necessary step in the activation of its natural tyrosine kinase receptor, c-kit [74]. The SCF/c-kit interaction leads to the first step of a biochemical cascade responsible for several effects, including cell proliferation. The first identified natural compound, named NSC 521777, which passes Lipinski's rules, is a cucurbitacin (Fig. 2.11), and it has been reported as an active anticancer compound involved in the tyrosine kinase signaling pathway [74]. The second one, named NSC 293161, is the natural antibiotic actinorhodin (Fig. 2.11) [74] and has been shown to possess a bis anthraquinonic structure. Biological assays of the top scored compounds are in progress.

Another example of PPI inhibition by small natural compounds has been described by Reindl et al. [75]. In this work, a screen of diverse chemical libraries, consisting of 22,461 small molecules, was carried out using a fluorescence polarization assay to identify compounds that could interfere with Polo-like kinase 1 (Plk1). Plk1 is a regulator of multiple stages of mitotic progression, and it is overexpressed in many types of human cancers. Its inhibition by small molecules has been demonstrated to induce mitotic arrest and apoptosis in cancer cells, both in vitro and in vivo [75]. Usually, inhibitors of Plk1 target the conserved ATP-binding site. The authors demonstrated that Plk1 can alternatively be targeted by small molecules that inhibit the function of the polo-box domain (PBD), a recently discovered protein domain that mediates the intracellular localization of Plk1 [75]. In particular, the natural compound thymoquinone (Fig. 2.11), the bioactive constituent of the volatile oil of black seed (*Nigella sativa*) [76], represents the first non-peptidic inhibitors of the PPI mediated by Plk1/PBD (apparent $IC_{50} = 1.14 \pm 0.04$ μM)

[75]. This compound and its synthetic derivative inhibit the function of Plk1/PBD in vitro and cause Plk1 mislocalization, defective chromosome congression, mitotic arrest, and apoptosis in HeLa cancer cells [75]. In this way, the anticancer activity of thymoquinone has been demonstrated, making it one of the few natural compounds found capable of inhibiting a PPI.

Another important target in the field of PPI inhibition is represented by the interaction between the Mcl-1 and Bak proteins [77]. The overexpression of the Mcl-1 protein in cancer cells results in the sequestration of Bak, a key component in the regulation of normal cell apoptosis. An investigation on the ability of marine-derived small natural products to disrupt the interaction between Mcl-1 and Bak led to the isolation of 13 oxy-polyhalogenated diphenyl ethers (O-PHDEs) from the marine sponges *Dysidea granulosa* and *Dysidea herbacea*. These molecules exhibited an interesting activity in a preliminary FRET assay (IC_{50} values of 4.1 and 2.1 μg/mL, respectively) [77]. The isolated secondary metabolites derive from an association between the above-cited sponges and a symbiotic cyanobacterium, named *Oscillatoria spongeliae* [77]. Unfortunately, the bioassay-guided fractionation did not lead to the isolation of pure constituents with a higher activity than the crude extracts. In fact, only three of these compounds exhibited significant IC_{50} values <10 μg/mL, with only one compound, named 2-(2′,4′-dibromophenoxy)-3,5-dibromophenol (Fig. 2.11), showing an IC_{50} value of 2.1 μg/mL [77]. A synergistic effect, due to the combination of O-PHDEs, was believed to occur in the crude extracts.

Small molecules of natural origin have also demonstrated to be capable of controlling protein transport by interfering with PPI. An example is represented by the transport of proteins from the nucleus to the cytoplasm by the exportin CRM1, which recognizes cargo proteins through a leucine-rich nuclear export signal (NES) [78]. By analogy with a truncated analog of the anguinomycins, which are potent anticancer agents active in the picomolar range that belong to the leptomycin family of natural products, goniothalamin (Fig. 2.11) was discovered [79]. This compound is a member of a class of styryl lactones isolated from plants of the genus *Goniothalamus*. Goniothalamin has been reported to induce cytotoxicity in breast cancer cells, with an IC_{50} value of ca. 1.5 μM, causing growth arrest and apoptosis [79]. These effects have been attributed to the inhibition of nuclear export by this natural compound at concentrations above 500 nM, as demonstrated by the immunostaining of Rio2 kinase in HeLa cells [79], based on the disruption of the PPI between CRM1 and cargo proteins [79]. With the discovery of the mechanism of action of this natural product, the search for structurally simpler leptomycin analogs useful for anticancer therapy is of great interest.

Novel inhibitors of PPIs may also be applicable in regenerative medicine for the treatment for neuronal diseases, especially for the generation of new neurons after stroke [80]. To this end, a rapid in vitro HTS for identifying natural compounds able to inhibit Hes1 dimer formation has been described [80]. The inhibition of Hes1 dimerization results in the acceleration of basic-helix-loop-helix (bHLH) activator transcription, which may promote the differentiation of neural stem cells (NSCs) into neurons. The developed assay was based on the use of fluorophore-

labeled Hes1 and Hes1 immobilized on microplates [79]. In the presence of an inhibitor, the level of Hes1 dimerization was reduced, as detected by a corresponding reduction in fluorescence. With this system, six compounds able to inhibit Hes1 dimerization were identified from a library of natural products composed by secondary metabolites isolated by the authors [80]. Three of the identified active compounds are natural products from myxomycetes; one was isolated from *Jasminum grandiflorum* and two from *Saraca asoca*. The most active compound was lindbladione ($IC_{50} = 4.1$ μM) (Fig. 2.11), which was isolated from the myxomycetes *Lindbladia ubulina*. Furthermore, using a cell-based reporter gene assay, it has been demonstrated that two of the identified compounds, lindbladione and lycogaric acid (Fig. 2.11), the latter of which was isolated from the myxomycetes *Lycogala epidendrum*, inhibited the Hes1-mediated suppression of transcription in C3H10T1/2 cells.

2.10 Perspectives

A few protein–protein interaction inhibitors are nowadays entering in clinical trials. The first MDM2 inhibitor that entered clinical development is RG7112 from Hoffmann-La Roche (clinicaltrials.gov identifiers: NCT01164033, NCT01143740, NCT00623870, and NCT00559533). Four phase I clinical trials have been conducted to date in patients with advanced solid tumors, hematologic neoplasms, or liposarcomas prior to debulking surgery. Preliminary clinical data indicated that RG7112 appears to be well tolerated in patients and shows initial evidence of clinical activity and a mechanism of action consistent with targeting of the MDM2–p53 interaction. Other candidates are along the same line therefore it becomes clear that the possibility to develop protein–protein interaction inhibitors that are effective anticancer drugs with the predicted mechanism of action is a concrete perspective. The monitoring of the drug development process can be implemented assisting each step from the computational design to the lead optimization, cellular pharmacology, and animal testing with dedicated strategies focused on PPIs inhibitors mechanism. The monitoring process includes analytical tools and pharmacodynamic biomarkers that will ensure an higher and accelerated success rate.

References

1. Jones S, Thornton J (1996) Proc Natl Acad Sci USA 93:13
2. Cardinale D, Salo-Ahen OM, Ferrari S, Ponterini G, Cruciani G, Carosati E, Tochowicz AM, Mangani S, Wade RC, Costi MP (2010) Curr Med Chem 17:826
3. Thiel P, Kaiser M, Ottmann C (2012) Angew Chem Int Ed Engl 51:2012
4. Arkin MR, Wells JA (2004) Nature Rev Drug Disc 3:301

5. Kaladhar DS, Sai PC, Rao PVN, Krishnachaitanya A, Rao DG, Rao VV, Reddy ER, Kumar SV, Kumar DV (2012) Bioinformation 8:437
6. Matveev VV (2010) Theor Biol Med Model 7:19
7. García-Sáez AJ (2012) Cell Death Differ 19:1733
8. Liu Q, Leber B, Andrews DW (2012) Cell Cycle 11:3536
9. Huart AS, Maclaine NJ, Narayan V, Hupp TR (2012) PLoS One 7:e43391
10. Groner B, Weber A, Mack L (2012) Bioengineered 3:320
11. Roberts KE, Cushing PR, Boisguerin P, Madden DR, Donald BR (2012) PLoS Comput Biol 8:e1002477
12. De Simone A, Kitchen C, Kwan AH, Sunde M, Dobson CM, Frenkel D (2012) Proc Natl Acad Sci USA 109:6951
13. Robinson WH, Steinman L (2011) Nat Biotechnol 29:500
14. Dwane S, Kiely PA (2011) Bioengineered Bugs 2:247
15. Mullard A (2012) Nature Rev Drug Discov 11:173
16. De Las Rivas J, Prieto C (2012) Methods Mol Biol 910:279
17. Reynolds C, Damerell D, Jones S (2009) Bioinformatics (Oxford, England) 25:413
18. Metz A, Ciglia E, Gohlke H (2012) Curr Pharm Des 18:4630
19. Betzi S, Restouin A, Opi S, Arold ST, Parrot I, Guerlesquin F, Morelli X, Collette Y (2007) PNAS 104:19256
20. Glanzer JG, Liu s, Oakley GG (2011) Bioorg Med Chem 19:2589
21. Dinesh, Goswami A, Suresh PS, Thirunavukkarasu C, Weiergräber OH, Kumar MS (2011) Bioinformation 7:21
22. Higueruelo AP, Schreyer A, Bickerton GRJ, Pitt WR, Groom CR, Blundell TL (2009) Chem Biol Drug Des 74:457
23. Fry DC (2008) Curr Protein Pept Sci 9:240
24. Fry DC (2006) Biopolymers 84:535
25. Pagliaro L, Felding J, Audouze K, Nielsen SJ, Terry RB, Krog-Jensen C, Butcher S (2004) Curr Opin Chem Biol 8:442
26. Morelli X, Bourgeas R, Roche P (2011) Curr Opin Chem Biol 15:475
27. Bourgeas R, Basse M, Morelli X, Roche P (2010) PLoS One (www.plosone.org) 5:e9598
28. Bahadur RP, Zacharias M (2008) Cell Mol Life Sci: CMLS 65:1059
29. Morohashi Y, Kan T, Tominari Y, Fuwa H, Okamura Y, Watanabe N, Sato C, Natsugari H, Fukuyama T, Iwatsubo T, Tomita T (2006) J Biol Chem 281:14670
30. Pennington RP, Wei Z, Rui L, Doig JA, Graham B, Kuski K, Gabriel GG, Mousseau DD (2011) J Neural Transm 118:987
31. Shearman MS, Beher D, Clarke EE, Lewis HD, Harrison T, Hunt P, Nadin A, Smith AL, Stevenson G, Castro JL (2000) Biochem 39:8698
32. Esler WP, Kimberly WT, Ostaszewski BL, Diehl TS, Moore CL, Tsai J-Y, Rahmati T, Xia W, Selkoe DJ, Wolfe MS (2000) Nature Cell Biol 2:428
33. Das C, Berezovska O, Diehl TS, Genet C, Buldyrev I, Tsai J-Y, Hyman BT, Wolfe MS (2003) J Am Chem Soc 125:11794
34. Dovey HF, John V, Anderson JP, Chen LZ, de Saint Andrieu P, Fang LY, Freedman SB, Folmer B, Goldbach E, Holsztynska EJ, Hu KL, Johnson-Wood KL, Kennedy SL, Kholodenko D, Knops JE, Latimer LH, Lee M, Liao Z, Lieberburg IM, Motter RN, Mutter LC, Nietz J, Quinn KP, Sacchi KL, Seubert PA, Shopp GM, Thorsett ED, Tung JS, Wu J, Yang S, Yin CT, Schenk DB (2001) Neurochemistry 76:173
35. Churcher I, Williams S, Kerrad S, Harrison T, Castro JL, Shearman MS, Lewis HD, Clarke EE, Wrigley JDJ, Beher D, Tang YS, Liu W (2003) J Med Chem 46:2275
36. Tian G, Ghanekar SV, Aharony D, Shenvi AB, Jacobs RT, Liu X, Greenberg BD (2003) J Biol Chem 278:28968
37. Kornilova AY, Bihel F, Das C, Wolfe MS (2005) PNAS 102:3230
38. Fan L-Y, Chiu M-J (2010) Acta Neurol Taiwan 19:228
39. Potter PE (2010) Dis J Am Osteopath Assoc 110(9 suppl 8):S27
40. Wolfe MS (2008) Neurotherapeutics 5:391

41. Wolfe MS (2008) Curr Alzheimer Res 5:158
42. Moore CD, Wu1 H, Bolanos B, Bergqvist S, Brooun A, Pauly T, Nowlin D (2009) Chem Biol Drug Des 74:212
43. Crisóstomo FRP, Feng Y, Zhu X, Welsh K, An J, Reed JC, Huang Z (2009) Bioorg Med Chem Lett 19:6413
44. Schimmer AD, Dalili S, Batey RA, Ried SJ (2006) Cell Death Differ 13:179
45. Sun H, Nikolovska-Coleska Z, Yang C-Y, Qian D, Lu J, Qiu S, Bai L, Peng Y, Cai Q, Wang S (2008) Chem Res 41:1264
46. Huang J-W, Zhang Z, Wu B, Cellitti J-F, Zhang X, Dahl R, Shiau C-W, Welsh K, Emdadi A, Stebbins JL, Reed JC, Pellecchia M (2008) J Med Chem 51:7111
47. Nikolovska-Coleska Z, Xu L, Hu Z, Tomita Y, Li P, Roller PP, Wang R, Fang X, Guo R, Zhang M, Lippman ME, Yang D, Wang S (2004) J Med Chem 47:2430
48. Weiss B, Bollag G, Shannon K (1999) Am J Med Genet (Semin Med Genet) 89:14
49. Goustin AS, Leof EB, Shipley GD, Moses HL (1986) Cancer Res 46:1015
50. Woods SA, Marmor E, Feldkamp M, Lau N, Apicelli AJ, Boss G, Gutmann DH, Guha A (2002) J Neurosurg 97:627
51. Vigi D, Cherfils J, Rossman KL, Der CJ (2010) Nat Rev Cancer 10:842
52. Cohen LH, Pieterman E, van Leeuwen REW, Overhand M, Burm BEA, van der Marel GA, van Boom JH (2000) Biochem Pharmacol 60:1061
53. Karnoub AE, Weinberg RA (2008) Nature Rev Mol Cell Biol 9:517
54. Danenberg PV, Malli H, Swenson S (1999) Semin Oncol 26:621
55. Carreras CW, Santi DV (1995) Annu Rev Biochem 64:721
56. Costi PM, Rinaldi M, Tondi D, Pecorari P, Barlocco D, Ghelli S, Stroud RM, Santi DV, Stout TJ, Musiu C, Marangiu EM, Pani A, Congiu D, Loi GA, La Colla P (1999) J Med Chem 42:2112
57. Stout TJ, Tondi D, Rinaldi M, Barlocco D, Pecorari P, Santi DV, Kuntz ID, Stroud RM, Shoichet BK,Costi MP (1999) Biochemistry 38:1607
58. Chu E, Callender MA, Farrell MP, Schmitz JC (2003) Cancer Chemother Pharmacol 52(Suppl 1):S80
59. Chu E, Voeller D, Koeller DM, Drake JC, Takimoto CH, Maley GF, Maley F, Allegra CJ (1993) Proc Natl Acad Sci USA 90:517
60. Scanlon KJ, Kashani-Sabet M (1988) Proc Natl Acad Sci USA 85:650
61. Lovelace LL, Gibson LM, Lebioda L (2007) Biochemistry 46:2823
62. Schiffer CA, Davisson VJ, Santi DV, Stroud RM (1991) J Mol Biol 219:161
63. Lovelace LL, Minor W, Lebioda L (2005) Acta Crystallogr D Biol Crystallogr 61(Pt 5):622
64. Gibson LM, Lovelace LL, Lebioda L (2008) Biochemistry 47:4636
65. Cardinale D, Guaitoli G, Tondi D, Luciani R, Henrich S, Salo-Ahen OM, Ferrari S, Marverti G, Guerrieri D, Ligabue A, Frassineti C, Pozzi C, Mangani S, Fessas D, Guerrini R, Ponterini G, Wade RC, Costi MP (2011) Proc Natl Acad Sci USA 108:E542
66. Lepourcelet M, Chen Y-N P, France DS, Wang H, Crews P, Petersen F, Bruseo C, Wood AW, Shivdasani RA (2004) Cancer cell 5:91
67. Doghman M, Cazareth J, Lalli E (2008) J Clin Endocrinol Metab 93:3222
68. Wei W, Chua MS, Grepper S, So S (2010) Int J Cancer 126:2426
69. Tsukamoto S, Yoshida T, Hosono H, Ohta T, Yokosawa H (2006) Bioorg Med Chem Lett 16:69
70. Schneekloth JS, Crews CM (2011) Curr Drug Targets 12:1581
71. Duncan SJ, Grüschow S, Williams DH, McNicholas C, Purewal R, Hajek M, Gerlitz M, Martin S, Wrigley SK, Moore M (2001) J Am Chem Soc 123:554
72. Hashimoto J, Watanabe T, Seki T, Karasawa S, Izumikawa M, Seki T, Iemura S, Natsume T, Nomura N, Goshima N, Miyawaki A, Takagi M, Shin-Ya K (2009) J Biomol Screen 14:970
73. Izumikawa M, Hashimoto J, Hirokawa T, Sugimoto S, Kato T, Takagi M, Shin-Ya K (2010) J Nat Prod 73:628
74. Alcaro S, Gontrani L, Incani O, Ortuso F (2008) Theor Chem Acc 120:523
75. Reindl W, Yuan J, Krämer A, Strebhardt K, Berg T (2008) Chem Biol 15:459

76. Gali-Muhtasib H, Roessner A, Schneider-Stock R (2006) Int J Biochem Cell Biol 38:1249
77. Calcul L, Chow R, Oliver AG, Tenney K, White KN, Wood AW, Fiorilla C, Crews P (2009) J Nat Prod 72:443
78. Gademann K (2011) Curr Drug Targets 12:1574
79. Wach JY, Güttinger S, Kutay U, Gademann K (2010) Bioorg Med Chem Lett 20:2843
80. Arai MA, Masada A, Ohtsuka T, Kageyama R, Ishibashi M (2009) Bioorg Med Chem Lett 19:5778

Chapter 3
Disrupting Protein–Protein Interfaces Using GRID Molecular Interaction Fields

Simon Cross, Massimo Baroni, Francesco Ortuso, Stefano Alcaro and Gabriele Cruciani

3.1 Introduction

Protein–protein interactions (PPI) are central to most biological process and include aspects such as viral self-assembly, cell proliferation, growth, differentiation, signal transduction, and programmed cell death [1], and the last decade has seen an explosion of interest [2–4]. The human interactome has been estimated to involve ~25,000 proteins and ~650,000 interactions [5], and currently, only about 0.3 % of these have been identified [6]. In silico approaches targeting PPIs are still limited, however the versatility of GRID Molecular Interaction Fields is demonstrated through several case studies that explore how novel PPI inhibitors can be identified. PPI therefore represent a large and important target for therapeutics, illustrated to some extent by therapeutic antibodies as a rapidly expanding segment of the drug market [7], an example of which is the monoclonal antibody PPI inhibitor Herceptin [8]. HIV viral fusion requires a series of PPIs to occur [9], the apoptosis pathway involving the Bcl2 family relies on several PPIs [10], misfunctions in a range of cancers, and HDM2 negatively regulates the transcriptional activity of p53 through a PPI [11]. There is of course a disadvantage to therapeutic antibodies, which is their higher cost of manufacture, lack of oral bioavailability, and non-cell permeable nature [7]. Up until about a decade ago, it was generally thought to be far too difficult to find small molecule therapeutics to

S. Cross (✉)
Molecular Discovery Limited, 215 Marsh Road, Pinner, Middlesex, London HA5 5NE, UK
e-mail: simon@moldiscovery.com

M. Baroni · G. Cruciani
Laboratory for Chemometrics and Cheminformatics, Chemistry Department, University of Perugia, Via Elce di sotto 10, I-06123 Perugia, Italy

F. Ortuso · S. Alcaro
Laboratory of Computational Medicinal Chemistry, Department of "Scienze della Salute", University "Magna Græcia" of Catanzaro, Viale Europa, Loc, Germaneto 88100 Catanzaro, Italy

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4_3,

inhibit PPIs [12] that the affinities of small molecules could not compete with the protein partner unless the molecules were too large to be 'drug-like'. Another potential problem is the large 'flat' nature of the PPI, with the contact interfaces reported as between 750–3000 $Å^2$ [2, 3, 7], whereas small molecule drugs and the associated design methodologies have typically targeted well-defined enzyme pockets (with a contact interface of 300–1000 $Å^2$) [3], and it has been reported that High-Throughput Screening (HTS) does not routinely find PPI inhibitors [13, 14]. Also, given that the binding partner is a protein, there are no natural small molecule binders to use as a starting point in the design of compounds to inhibit the interaction. Clearly, there are some different challenges in finding PPI inhibitors than in conventional drug discovery, but in recent years, there has been some progress in the area, providing cause for optimism in targeting such a potentially valuable area. One finding is that a smaller subset of residues at the contact interface contributes most to the free energy of binding that these 'hotspots' constitute less than half of the contact interface and are usually found at the centre of the interaction. Six examples of small molecule PPI inhibitors are reviewed by Wells & McLendon, which are interleukin-2 (IL2) and its receptor, B-cell lymphoma-2 family members and the pro-apoptotic molecules BAK/BAD, human protein double minute 2 and tumour suppressor p53 human double minute 2, p53 (HDM2-p53), human papilloma virus transcription factor E2 human papilloma virus (HPV-E2) and helicase E1, bacterial FtsZ and ZipA, and the cytokine tumour necrosis factor (TNF) and its receptors. For all six examples, the small molecule inhibitor binds to a pocket and was identified by screening, and for most examples, the affinity of binding is comparable to that of the native protein partner [3]. Further highlighting the potential optimism for disrupting PPI is the fact that for Bcl-X_L-BAK-BAD, Abbott's Navitoclax has currently demonstrated success in clinical trials [15].

Recently, the TIMBAL database for published PPI inhibitors has been reported by Higueruelo and co-workers [12], with small molecules disrupting 17 PPIs taken from 40 publications. Characterisation of the inhibitors reveals key differences compared with conventional drugs and screening molecules, with them tending to be relatively large, hydrophobic, contain more rings and fewer rotatable bonds, perhaps explaining the historical difficulties in finding such inhibitors by High-Throughput Screening (HTS).

In this chapter, we will briefly review current in silico approaches to target protein–protein interactions and present a methodology to identify the key protein–protein interface interactions using GRID molecular interaction fields (MIFs) [16], refined using the GRID-based pharmacophore model approach (GBPM) [17], and combined with the FLAP receptor-based pharmacophore screening method [18].

3.2 Computational Mapping of Protein–Protein Interactions

Mapping protein–protein interaction (PPI) networks is one of the current goals of proteomics, determining function from a protein's position in a complex web of protein–protein interactions. Experimental determination of PPI is possible through techniques such as co-immunoprecipitation, affinity chromatography, yeast two-hybrid, and mass spectrometry. Additionally, the genomic context of PPI can be used since interacting proteins are generally co-expressed [19–21]; for example, two proteins may be interacting if they are expressed at the same time and location. Even in the absence of structural or sequence information, one can detect the evolutionary fingerprints of pairs of interacting proteins from their genomic context [22].

Interaction sites tend to contain the most hydrophobic surface clusters [23–26] and analysing the residue composition can not only help to predict these sites, but also different types of protein–protein interaction [27]. Using more detailed parameters, such as solvation potential, residue interface potential, hydrophobicity, planarity, protrusion, and accessible surface area, has been shown to yield accurate predictions of the sites, although the prediction success rate was $\sim$66 % [28]. One computational method aimed at predicting interaction sites is protein–protein docking, and the performance of different methods is assessed by the community-wide blind test experiment CAPRI [29] (critical assessment of predicted interactions). Other methods to determine the interacting residues are based on the idea that these residues are evolutionary conserved and use the information from multiple sequence alignment to analyse this conservation, with a prediction success rate of $\sim$70 % [30–32].

Once the protein–protein pair is identified, it is important from a drug discovery perspective to be able to characterise the nature of the interaction in order to modulate it. Therefore, the availability of structural information about the interacting proteins is important; once the interaction site is characterised to find the 'hot spots' that are critical for binding this can provide information that aids the targeting of the site to inhibit the interaction. Experimental identification of these hot spots can be achieved using methods such as alanine scanning [33], and databases containing these experimentally determined hot spots are available (ASEdb [34], BID[35]); however, alanine scanning requires a large experimental effort that cannot be easily applied in a high-throughput manner, and hence, there is a need for computational methods to identify these hot spots.

Hot spot characterisation has been performed computationally[36, 37] showing that they are enriched in Tyr, Trp, and Arg residues and surrounded by energetically less important residues whose role is to occlude bulk water[38]. Leu, Ser, Thr, and Val residues are strongly disfavoured in hot spots, and Asn and Asp residues are more prevalent than Gln and Glu.

A natural progression from experimental method to in silico simulation has been demonstrated by in silico alanine scanning approaches based on the

estimation for each residue of the binding free energy difference upon mutation from wild type to alanine; residues strongly involved in binding are therefore highlighted with a large change in this difference [39–43]. Feature-based methods have also been applied to discriminate hot spot residues by using sequence and/or structural data [44–49]. One of the more recent is APIS [50], where 62 features from sequence and structure were investigated, redundant and irrelevant features were removed, and 9 individual feature-based predictors were developed to identify hot spots using an SVM approach. Another approach towards identifying hot spots is demonstrated by Vajda and co-workers [51–53] where the druggability of the site is an important consideration. The rationale is based upon the current trend in drug design involving experimentally finding hot spot interactions in proteins using fragment-based screening. These experimentally derived hot spots bind a variety of small molecules, and the number of different probe molecules binding to a site is predictive of its overall druggability. Vajda and co-workers use an in silico analogue of these experimental fragment-based approaches called computational solvent mapping [54]. Molecular probes are placed on a grid surrounding the protein, and favourable interaction regions are identified using empirical free energy functions. Conformations are clustered and ranked, and those regions that bind several clusters are identified as the hot spot regions. Additionally, accessible sidechain conformations of the important residues at the site are then varied to generate multiple alternative structures, then maps each alternative hot spot site conformation and selects the most likely conformation as that binds the highest number of clusters. The approach is not dissimilar to that used by the software GRID [16, 55], and it is this approach that we will focus on in detail in the remainder of this chapter.

3.2.1 GRID Molecular Interaction Fields

The GRID method, first published by Goodford in 1985 [16], is an attempt to compute interaction potentials with a target, which includes chemical specificity; over many years, the GRID forcefield was developed empirically by studying experimental observations. At the time of writing, there are 64 different chemical *probes*, each of which with different characteristics, that interact with a *target* molecule to produce molecular interaction fields (MIFs). A three-dimensional model of the target (for example, coming from experimental X-ray crystallographic data or a predicted conformation of a small molecule) is surrounded by an imaginary orthogonal grid, and the interaction potentials are calculated between the probe and the target at each position on the grid. Visualisation of these potentials is typically performed by plotting the isocontour using favourable interaction energies, allowing key interacting regions to be inspected. GRID is different from simply calculating the molecular electrostatic potentials around the target: the probes are anisometric, the target 'responds' when the probe is moved around it, and the assumption is made that both the target and probe are immersed

in water. For example, the carbonyl oxygen probe consists of an oxygen atom with 2 sp2 lone pairs; it has a size, polarizability, and electrostatic charge, and each lone pair can accept one hydrogen bond. The probe is placed at each grid point, and if no bad close contacts are found then nearby hydrogen bond donor atoms on the target are searched, and the probe is allowed to rotate to make the best possible hydrogen bond interactions with the target donor atoms, before finally computing the energy of interaction. The aromatic sp2 hydroxyl probe differs in several respects; the probe has a larger polarizability, makes hydrogen bonds with a different strength and can accept only one hydrogen bond, but can also donate one hydrogen bond. If both hydrogen bonds are made simultaneously, they are constrained towards the sp2 angle of 120°. Other probes differ again with respect to these parameters, and therefore, the chemical specificity is obtained. The target also responds to the probe; in addition to rotating the probe to make the best hydrogen-bonding interactions, the target hydrogen atoms are by default also rotated in the search. With the more advanced MOVE directive, entire sidechains are allowed to move in response to the probe, again to make the most favourable interactions possible. With the assumption that the probe and target are both immersed in water, the most reasonable assumption for biology, a bulk dielectric constant of 80 is used that diminishes towards 4 in the deep centre of a large globular macromolecule.

MIFs therefore provide information about where the favourable and unfavourable locations for the probe around the target are located. The energy function describing this was chosen to represent the underlying physical interactions and, as described above, was parameterized against experimental observations. The function takes the following form:

$$E = E_{\mathrm{VDW}} + E_{\mathrm{EL}} + E_{\mathrm{HB}} + S$$

where E_{VDW} is the van de Waals energy, E_{EL} is the electrostatic energy, E_{HB} is the hydrogen bond energy, and S is an entropic term; a more detailed description of these terms in GRID can be found here [16, 56, 57]. The entropic term, in particular for the hydrophobic probe, known as the DRY probe, is worth describing in more detail. The favourable entropic contribution due to the displacement of a single water molecule from a hydrophobic surface is assumed constant. The value is calculated by comparing the possible hydrogen bond combinations that the water molecule can form when at the hydrophobic surface and when in bulk water. In bulk water, it is assumed that three of a possible four hydrogen bonds are formed, with four combinations of these, therefore giving $RT\ln(4) = -0.848$ kcal mol^{-1}, which is larger than a typical attractive Lennard-Jones interaction (~ -0.2 kcal mol^{-1}), but smaller than a typical hydrogen-bonding interaction (-2 to -4 kcal mol^{-1}). The DRY probe therefore includes this entropic term; for the enthalpic contributions, it makes Lennard–Jones interactions in the same way as a methyl probe, is neutral, and has no electrostatic interactions. The hydrogen-bonding interactions of a water probe at the same position are subtracted from the DRY enthalpic term, reflecting the fact that the hydrophobic probe is

unable to make these interactions at polar parts of the target and will be disfavoured at these positions. The DRY probe is very useful for detecting hydrophobic patches on both proteins and small molecules, and is therefore used routinely across the many applications of GRID [58].

Describing how a target molecule appears at a molecular level to an external chemical observer is incredibly useful for the primary application area of drug discovery, and the original application of GRID was to map protein binding sites for structure-based design. Contouring the binding site MIFs for different chemical probes can aid greatly in the design of new small molecule inhibitors, and this was achieved by von Itzstein in one of the first examples of rational drug design, where GRID MIFs were used in the design of Zanamivir by identifying a favourable region for the guanidino group [59]. From a small molecule perspective, MIFs for inhibitors aligned in the same frame of reference can be used comparatively in combination with chemometric methods to identify interacting regions that are more important in terms of their contribution to the inhibitor's activity; the 3D-QSAR (quantitative structure–activity relationship) approach. Comparing the MIFs for a set of aligned protein targets in an analogous approach can also allow the optimisation of selectivity between the targets [60].

Understanding of the pharmacokinetics of a drug can be achieved by analysing its water (OH2 probe) and DRY MIFs (simulating the interactions of the molecule with the aqueous phase and lipid membrane), in addition to various descriptors derived from these MIFs, as in the VolSurf approach [61]. In the area of predictive metabolism, mapping the various Cytochrome P450 isoforms using GRID MIFs, and comparing these MIFs with those of small molecule substrates, in combination with chemical reactivity prediction, enables the prediction of metabolic 'soft spots' and the optimisation of metabolic stability, as in the MetaSite approach [62]. Descriptors calculated from small molecule MIFs using a circular fingerprint approach have also proved useful in predicting the pKa values of titratable groups, as in MoKa [63]. Fragment or scaffold 'hopping' is also possible by comparing new fragment MIFs to those of a ligand template or those from a target binding site, which is an invaluable aid when trying to jump into novel chemical space or modulate pharmacokinetics by finding a new chemical moiety containing the interactions essential for activity but differing in other locations [64].

One of the key benefits of GRID MIFs in recent years has been in combination with pharmacophore quadruplet fingerprints, allowing the comparison of small molecules, protein receptors, and combinations of both using a common reference framework, as in the FLAP approach [18] which is described in more detail in the next section.

3.2.2 FLAP

The background to GRID molecular interaction fields and their widespread application has been described above; however, one of the limitations has been

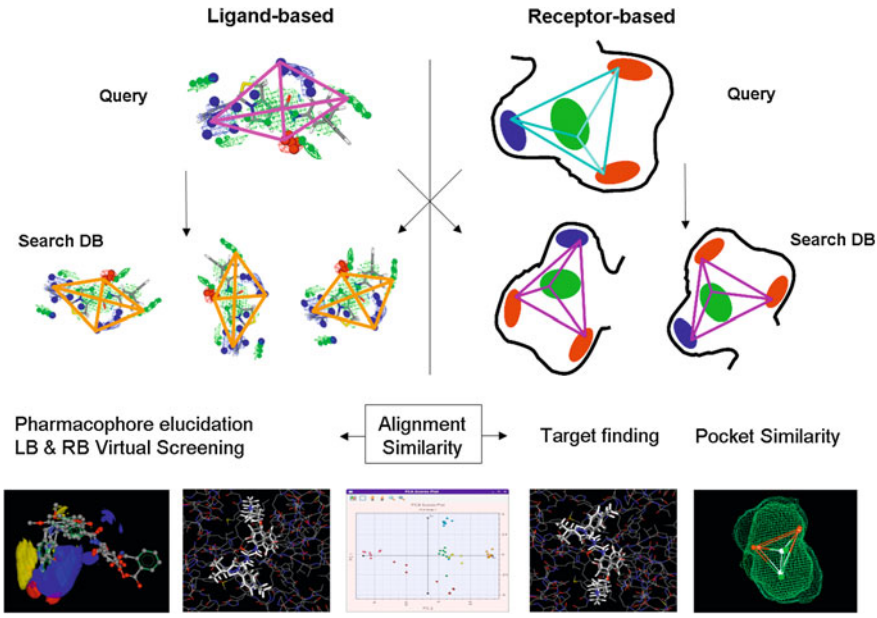


Fig. 3.1 Schematic view illustrating how GRID molecular interaction fields are used along with quadruplets formed from representative minima points to allow molecular alignment and field similarity scoring in a common reference framework using FLAP

aligning the various molecules correctly to compare their MIFs. For example, in a structure-based design approach, it is desirable to locate a potential lead molecule inside the receptor-binding site in an optimal way to understand if it can form favourable interactions. For 3D-QSAR, it is essential to align the lead series correctly to highlight small MIF differences that can lead to large changes in activity. When building a potential pharmacophore model for a group of compounds, it is also important to test a variety of different possible alignments to gain an understanding of whether there is a common pharmacophoric interaction field that can then be used to compare with new test molecules to find new leads or gain an understanding of the SAR. When investigating selectivity between targets, it is useful to have the different sites aligned to visualise the differences between the MIFs to aid design of a ligand to match one but not the other.

FLAP was designed to overcome these limitations, and given that small molecules and protein targets can both be described in terms of the MIFs, there is a common reference framework (see Fig. 3.1) that allows the comparison of small molecule to small molecule (for example, ligand alignment and virtual screening), small molecule to receptor (pose prediction in an analogous sense to docking and virtual screening), receptor to receptor (binding site comparison for selectivity, prediction of off-target effects), and ligand and receptor-based pharmacophore elucidation (for understanding SAR, and as a template for virtual screening). The FLAP method essentially consists of two aspects: the alignment step and the scoring

step. For each molecule (and each predicted conformer) being considered, the GRID MIFs are calculated. Typically, these are provided by the H, O, N1, and DRY probes which describe the shape, hydrogen bond acceptor, hydrogen bond donor, and hydrophobic interactions, respectively. For the O, N1, and DRY MIFs, the most relevant discrete points are selected using a weighted energy and spatial distribution. The points are combined to form a large number of quadruplets. Similar quadruplets between two different structures are then used for alignment. Once the alignment is performed, the similarity between the MIFs of the molecules is calculated, so for example, the relative hydrophobic similarity can be compared, or the global similarity or the shape similarity, etc. The most similar conformations and alignments are then selected to represent the overall alignment score for the structure.

When comparing the MIFs between small molecules and receptors, the receptor MIFs will match where the atoms in the small molecule that are similar to the MIF probe should be located, and vice versa. For this reason, pseudoMIFs are also generated that describe the potentials around an atom in an analogous fashion to an electron density map, and these can then be directly compared to the MIFs from the interacting partner. Since the probes used to generate the MIFs are allowed to rotate to form the most optimal interaction, whereas a placed ligand's atom is not, the similarity is computed in both directions, and the overall similarity is obtained from the product of each. For example, this ensures that a ligand hydrogen bond donor atom located well inside the corresponding receptor MIF is oriented correctly, since its corresponding hydrogen bond acceptor MIF must contain the receptor acceptor atom.

In addition to the alignment and similarity of two structures using their respective quadruplets and MIFs, the quadruplets can also be stored as a binary fingerprint, enabling alignment-independent comparison for the much higher throughput required by large scale virtual screening.

The utility of FLAP for virtual screening has been demonstrated many times on internal projects and can be highlighted in the following prospective and retrospective cases. Using the structure of a known calcium channel antagonist, a virtual screening cascade was performed which included FLAP as a ligand-based pharmacophoric similarity filter. Six compounds were obtained from a final pool of 20 tested compounds which demonstrated an ionotropic potency of the same magnitude as the known antagonist Diltiazem, including three novel chemotypes [65]. In the retrospective study, FLAP was applied to the DUD benchmark dataset [66] to recover known actives from selected decoy structures. The dataset contains 40 targets; for each target, a set of decoy compounds was carefully selected such that they contained similar molecular properties to the known actives molecules, with the aim of excluding trivial bias. The primary limitation of the DUD dataset is the number of structure analogues present for some of the targets; therefore, for the FLAP validation, we focused on chemotype retrieval using 13 targets for which there were 15 or more chemotypes [67]. A number of different approaches were tested, using single template ligand-based and receptor-based approaches, and also several data fusion approaches. For the single template approaches, at a false

positive rate of 1 %, typically 20–30 % of the chemotypes were found. For the data fusion approaches, this figure rose to ~40 % recovery.

FLAP has therefore been demonstrated to be extremely useful when aligning molecules and scoring them according to their MIF similarity. This aspect prompted us to extend the method to pharmacophore elucidation from a set of known active molecules with the question: rather than finding common rule-based pharmacophoric features and then aligning the molecules based on these, can we find a set of common alignments (ideally the bioactive conformations) using their MIFs and extract the common pharmacophore from this alignment set? Given a set of aligned structures, the MIFs can easily be reduced to represent the common pharmacophoric interaction field by taking the MIF average across the dataset. The advantage of this is that the interacting features do not have to be common to a user-specified number of molecules in advance; fields regions common to 5 out of 6 molecules will simply be 5/6th the intensity; therefore, partial matching is inherently taken into account. The same is true of the pseudoMIFs, and ultimately, a pseudomolecule is produced, analogous to any other molecule that can be used for alignment and/or virtual screening. To generate the common alignments, a conformer selection cascade is performed to reduce the search space from 100–1,000 s of conformers per molecule to ~10; a pruned tree search is then performed testing each molecule/conformer as the template, and sequentially aligning the other molecule conformers to this, scoring the field similarity across the alignment set, and finally selecting the best alignment models. Validation of the common alignments using the dataset from Patel et al. [68] has shown very promising results in terms of reproducing the experimental bioactive conformation [69]. Testing the discrimination using the DUD dataset as outlined above, and building models using 2D structures of the known actives before screening with the pharmacophoric pseudomolecule returned ~30 % of the chemotypes with only 1 % of the false positives, a return better than any of the other single template methods, including the receptor-based pharmacophore filtered by the X-ray bound ligand. Since the alignment and discriminatory power of the FLAP approaches has shown to be so promising, we decided to investigate the application to disrupting protein–protein interactions; as described earlier, finding small molecules to disrupt these interactions is of considerable potential therapeutic benefit. Using GRID MIFs to describe the key protein–protein interactions and applying the FLAP approach to use these key interactions in a receptor-based pharmacophore approach to find novel inhibitors are a logical extension and another demonstration of the power of using a ‘common reference framework’.

3.2.3 GBPM and FLAP

The GRID-based pharmacophore model (GBPM) was developed in one of our labs (Ortuso/Alcaro), with the aim of generating pharmacphore models useful for QSAR and virtual screening experiments by means of an unbiased computational

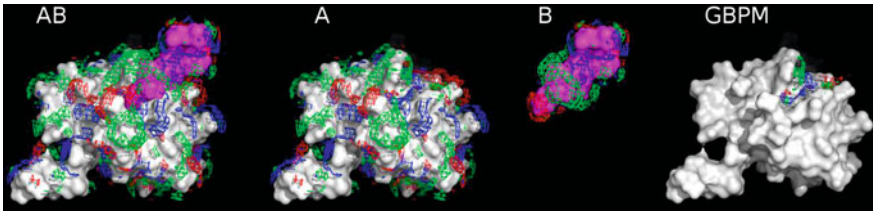


Fig. 3.2 The $\alpha + \beta$ complex (*AB*), α subunit (*A*), and the β subunit (*B*) are combined using logical operations to reveal the GRID molecular interaction fields focused at the interaction site

protocol [17, 70], using GRID MIFs to determine the pharmacophoric interactions at the interface of a generic complex.

The GBPM method is relatively straightforward and follows a six-step procedure. The first step is a pre-treatment step, preparing from a PDB file of the complex separate files of the complex (subunits $\alpha + \beta$), the receptor (subunit α), and the ligand (subunit β). The ligand can be a generic molecule (i.e. another protein, nucleic acid, or small molecule); therefore, for protein–protein interactions, it represents the partner that the desired small molecule ligand antagonist should 'mimic'. The second step involves the calculation of the GRID MIFs for each of the three targets from step 1. In the third step, the GRID GRAB utility is used, whereby an energy comparison is performed between the MIFs, generating new maps where the fields are focused on the interaction areas. The MIFs for subunit α (the receptor) contain the information for this interaction area, in addition to extra regions that are not involved in the binding of the complex. These extra-field regions can be found from the subunit $\alpha + \beta$ (the complex), which also contains regions from the exposed β subunit. Therefore, by performing logical operations on the three subunits, the MIFs for the α subunit describing only the interaction area with the β subunit can be extracted, and without any user bias (see Fig. 3.2).

The fourth step involves extracting relevant points from the MIFs; therefore, typically the probes DRY, O, and N1 are used, describing the hydrophobic, hydrogen bond acceptor, and hydrogen bond donor interactions, respectively. More sophisticated models can also be built by including other probes (for example halogen or charged probes). For each of the probes, MIF points are extracted that are typically 5–15 % higher than the global minimum value for that probe, representing the most intense interactions. For a typical feature-based pharmacophore model, it can be the case that this results in the selection of too many points, making the query too complex to be useful. In this case, the thresholding can be modified to simplify the query. The fifth step in the procedure simply involves merging each of the selected features into a single pharmacophoric model.

The sixth step involves the validation of these points as a model and modulating the number of features. The quality of the model is tested in terms of recognition of the original ligand (subunit β), typically by using the CiTest fit module in the software Catalyst [71], which calculates a non-energy weighted fit value; the

pharmacophoric points are converted into Catalyst features, and the GRID energies are included according to the following equation:

$$wF_{ij} = EF_i/AEF_j$$

where wF_{ij} is the weight for the feature i into the hypothesis j, EF_i is the GRID energy for the features i, and AEF_j is the average GRID energy value computed along all probes of the hypothesis j. The approach results in a maximum fit value (MFV) equal to the total number of features available for each hypothesis. Typically, many models can be produced from the GBPM procedure, and all of these are submitted to this validation to identify the best one. A fit index (FI) is defined as the ratio between the CiTest fit, and the MFV is used for the evaluation and identification of the best GBPM, since it also allows the comparison between models with different numbers of features. This FI descriptor can also be used to test other known ligands that bind at the same site to validate the model.

As described above, the software FLAP is able to perform molecular alignment and pharmacophoric similarity based on the GRID fields, without the need to subjectively extract and convert MIF minima point 'hotspots' into simplified rule-based features that are required by classical pharmacophore methods. The GBPM approach has been adapted and incorporated into FLAP in a straightforward manner. GBPM step 1 is performed as described above; then, steps 2 and 3 are performed automatically inside FLAP. Instead of extracting a few hot spots to derive the rule-based features in steps 4 and 5, a large number of points are extracted from each MIF and these are used to define the FLAP quadruplets. These quadruplets can be used in the standard FLAP manner to search for corresponding quadruplets in small molecule ligands; once the alignment is achieved, the MIF similarity between the ligand and the GBPM-receptor-interaction site MIFs is calculated. The validation is simply performed by analysing the similarity of the original ligand, and any known test ligands, to these receptor MIFs. Known decoys can be used to judge the discrimination of the model; then, if discrimination is observed, databases of commercially available compounds can be screened in the usual fashion.

It is worth noting that the GBPM fields are describing the interactions on the target at the interface where the ligand binds; an advantage of this is that they are not only focused on the common pharmacophoric interactions between the ligand and target, but also include regions where the ligand may not be making the optimum interactions, providing scope to find alternative ligands with improved interactions. The FLAP-GBPM approach also reaffirms the power of FLAP to provide a 'common reference framework'—alignment and MIF similarity are used in the same fashion as for small molecule ligand-based alignment or receptor-binding site pose prediction or binding site-binding site comparison; the only difference is the manner in which the relevant MIF regions are extracted. In the next section, we will discuss some examples applying FLAP-GBPM to disrupting protein–protein interactions.

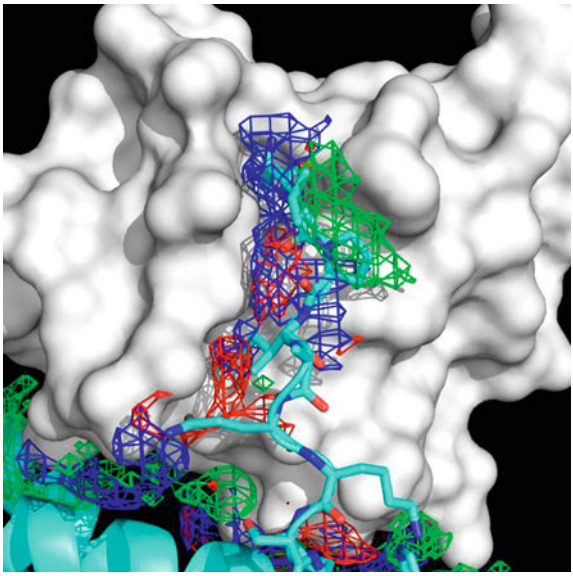


Fig. 3.3 Smac (*cyan*) bound to the BIR3 domain of XIAP (*white surface*), with the GBPM-derived GRID molecular interaction fields (hydrophobic = *green*, hydrogen bond acceptor = *red*, hydrogen bond donor = *blue*)

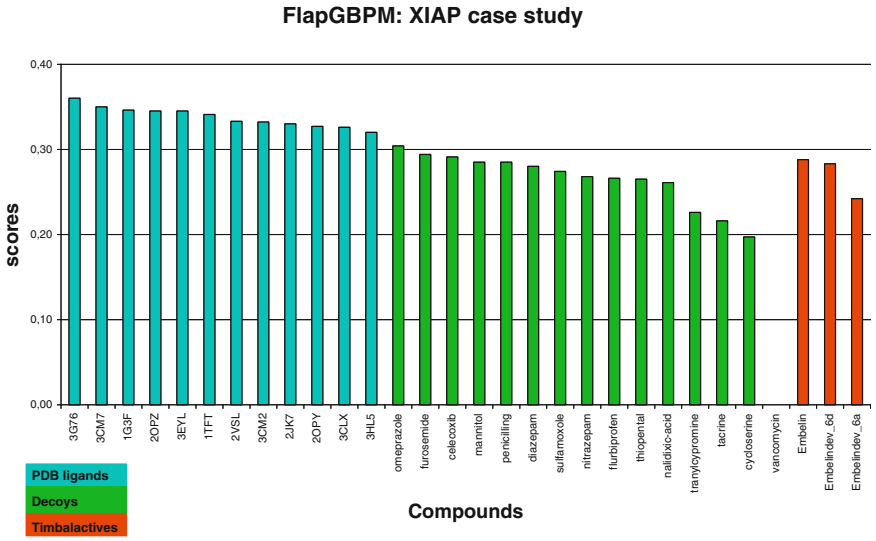
3.2.4 Case Studies

3.2.4.1 X-Linked Inhibitor of Apoptosis

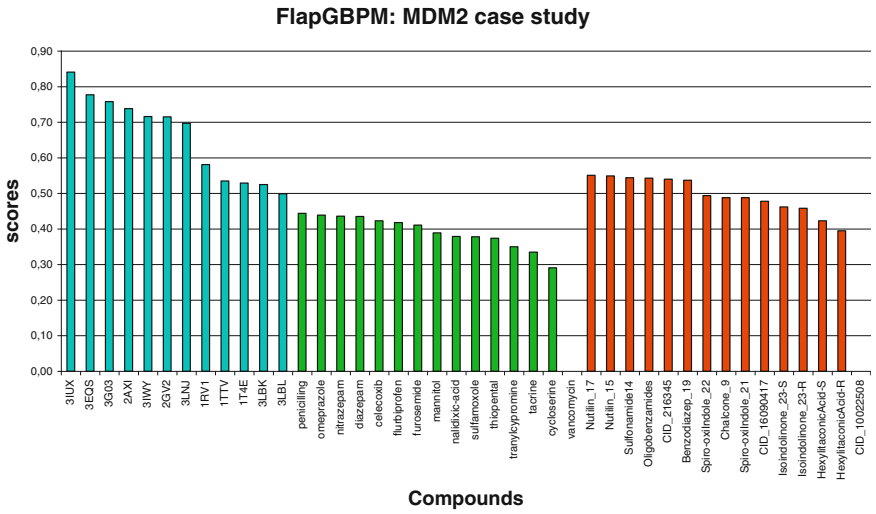
This target is a member of a family of proteins involved in the regulation of apoptosis by inhibiting the Caspase family of enzymes. A mammalian protein called Smac binds to the IAP proteins preventing the Caspase inactivation; two structures of the third baculovirus IAP repeat (BIR3) domain of XIAP complexed to Smac and to Caspase-9 have been published (PDB IDs: 1G73 [72] and 1NW9 [73], respectively). Thirteen other active ligands are also available in PDB entries 1TFT and 1TFQ [74], 1G3F [75], 2JK7 [76], (2OPY, 2OPZ) [77], 2VSL [78], 3EYL [79], (3CLX, 3CM2, 3CM7) [80], 3G76 [81], 3HL5 [82]. The XIAP-Caspase recognition area is about 700 $Å^2$, involving several hydrophobic, electrostatic, and hydrogen-bonding interactions; therefore, a traditional receptor-based pharmacophore model is relatively difficult to derive. The low number of molecules recognising the XIAP BIR3 domain also means that a traditional ligand-based pharmacophore elucidation approach is also difficult to achieve, and for these reasons, GBPM was applied in the original study by Ortuso, Langer, and Alcaro [17]. Here, we have extended the previous work, including a larger number of ligands and using the FLAP-GBPM approach which allows derivation of the critical interaction MIFs and the direct use of these focused MIFs to screen other ligands.

The complex from 1G73 was used to derive the FLAP-GBPM model, with the interaction MIFs shown in Fig. 3.3.

Graph 3.1 XIAP case study: PDB ligands, decoys, and Timbal active compounds screening results



Graph 3.2 MDM2 case study: PDB ligands, decoys and Timbal active compounds screening results



In the original study, the GRID program MINIM was used to identify low-energy points in the fields to use as Catalyst pharmacophore features: 4 features

from the N1 probe, 3 from the DRY probe, and 1 from the O probe were converted into Catalyst features as hydrogen bond donors, hydrophobic, and hydrogen bond acceptors, respectively. Given the nature of this 'complex' pharmacophore, the Catalyst fit index (FI) was relatively low, and so several rounds of removing features, re-scaling, and re-testing were performed until a good FI value was obtained. With FLAP-GBPM, this is not necessary, and even more potential user bias is removed from the procedure. The MIF similarity between the native ligand and the GBPM interaction MIFs is produced as a result of the analysis; when alternative ligands are screened, the equivalent similarity is obtained and can be directly compared to the original 'benchmark' value.

All previously reported PDB structures were downloaded, containing both protein–protein and protein–ligand interactions, and thirteen of these are used in the screening (currently, FLAP-GBPM is only able to screen small molecules as potential ligands). Three additional known ligands were obtained from the TIMBAL database [12], and a small 'decoy' dataset selected from fifteen known drugs as a proof-of-concept. The FLAP-GBPM method was applied to generate the interaction MIFs, and these used to align and score the decoy molecules (all small molecules were subjected to a conformation search, with the 30 lowest-energy conformers aligned to the GBPM MIFs and the best score chosen to represent the molecule).

The thirteen ligands from the first set of downloaded structures ranked higher than all other compounds, with the TIMBAL structures ranked amongst the top-scoring decoys (Graph 3.1). In our opinion, this demonstrates excellent discrimination, as has been shown previously with other applications of FLAP.

3.2.5 MDM2-p53

The p53 tumour suppressor protein can be bound by MDM2, increasing its degradation, and thus blocking p53's transcriptional activity that results in its tumour suppression activity [83]. Modulating the interaction of MDM2 with p53 is therefore of great interest in cancer treatment, and a range of molecules have been successfully produced that disrupt MDM2-p53 complexes (see 3 for a review). This protein–protein interface was not considered previously using GBPM; here, we decided to analyse whether a FLAP-GBPM model could be obtained from the structure of a p53 peptide interacting with MDM2 (PDB ID: 1T4F [84]) and how the interacting FLAP-GBPM MIFs were able to score the known ligands and several decoy ligands. Thirteen entries from the PDB (1RV1 [85], 1T4F, 1T4E [86], 1TTV [86], 2AXI [87], 2GV2 [88], 3EQS [89], 3G03 [90], 3IUX [91], 3IWY [92], 3LBK [93], 3LBL [95], 3LNJ [94]) were downloaded including both protein–protein and protein–ligand interactions; twelve of these contained small molecule ligands that were aligned and scored to the 1T4F GBPM model. In addition, fifteen other ligands were obtained from the TIMBAL database [12] and fifteen decoy molecules were selected from known drugs. All small molecules were subjected to the conformational searching described in the XIAP case above. Figure 3.4 shows

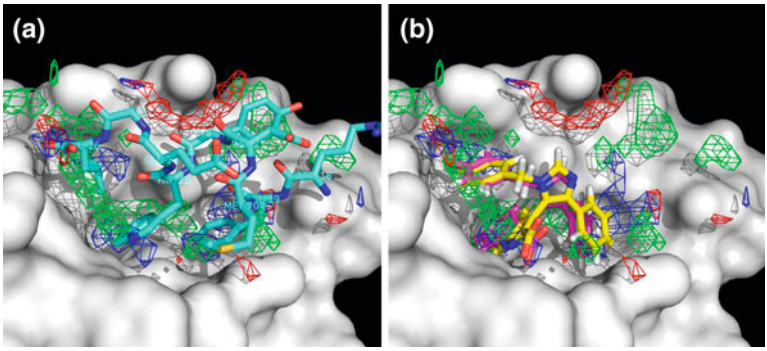


Fig. 3.4 The GBPM molecular interaction field maps focused on the interaction site between HDM2 and an optimized p53 peptide (*A*). The same interaction maps shown with the inhibitor from 3LBK (*B*) after protein alignment (*carbon atoms coloured yellow*) and the same inhibitor positioned by FLAP in the site for scoring (*carbon atoms coloured magenta*)

the GBPM interaction fields for MDM2-p53, in addition to a small molecule inhibitor from 3LBK.

The interaction fields show a broad hydrophobic character, fulfilled by the p53 isoleucine, tryptophan, tyrosine, and phenylalanine residues. The tryptophan residue also makes a hydrogen-bonding interaction with the MDM2 backbone carbonyl, and the terminal carboxylate on the isoleucine interacts with a tyrosine residue in MDM2. Additionally, the backbone carbonyl on the p53 peptide is in close proximity to the fields given by the GRID O probe, indicating a potential interaction with a MDM2 lysine residue. Figure 3.4 also illustrates the pose predicted by FLAP to best fulfil these interactions, which is positioned almost exactly in accordance with the experimental ligand position. In terms of scoring, twenty-five of the known inhibitors were clearly differentiated using their similarity to the interaction fields, with only two scoring lower than a number of the decoys (Graph 3.2).

In our opinion, this again represents excellent differentiation and illustrates the promise of this approach.

3.2.6 IL-2/IL-2Rα

Interleukin-2 (IL-2) is an immunoregulatory cytokine at the heart of the immune response, which binds to the alpha, beta, and gamma chains of its receptor (IL-2R). After the activation of T cell receptors by peptide major histocompatability complexes, IL-2 and IL-2Rα are expressed. The subsequent autocrine interaction of IL-2 with its receptors stimulates the proliferation of T cell, B cell, natural killer cell, and clonal expansion [95]. IL-2Rα is not expressed by resting T cells and B cells, but is continuously expressed by abnormal T cells in forms of leukaemia, autoimmunity, and transplant reject, and is therefore a potentially important therapeutic target to modulate in patients with these conditions. Since

crystallographic data are available for the structure of IL-2 complexed with its receptor, and additionally some small molecule inhibitors, this represents another interesting case study to investigate with the FLAP-GBPM approach.

Five PDB entries were downloaded including both protein–protein (PDB ID: 1Z92 [96]) and protein–ligand interactions (PDB IDs: 1M48 [97], 1M49 [97], 1PW6 [98], and 1PY2 [98]). The PDB model 1Z92, reporting the structure of IL-2 complexed with its receptor, was used as the input for the FLAP-GBPM method. Six known small molecules inhibitors were selected, four were obtained from the PDB models and two from the TIMBAL database [12], respectively. Additionally, as for the previous case studies, fifteen decoy molecules were aligned and scored using the approach to investigate the discriminatory ability of the identified interaction fields.

In this case study, FLAB-GBPM strongly demonstrated the capability to distinguish between active and inactive compounds. Only the six known inhibitors were recognized whilst all decoys were discarded. The graphical inspection of the predicted poses of the inhibitors revealed that all bind IL-2 in the same area, even if, for PDB compounds, the X-ray pose was not perfectly reproduced. The reason for this is likely due to induced fit phenomena, as suggested by the comparison of the PDB models. In 1Z92, the only protein–protein model available, the Phe42, Tyr45, and Leu72 sidechains display different conformations with respect to the other selected PDB models containing protein–ligand complexes; these residues appear to move to allow the small molecule ligands to bind more favourably. Since the FLAP-GBPM method uses the interaction maps from the static X-ray structure of the input protein, the map describing the shape of the receptor provides a steric constraint that prevents the X-ray poses from being reproduced (see Fig. 3.5). This

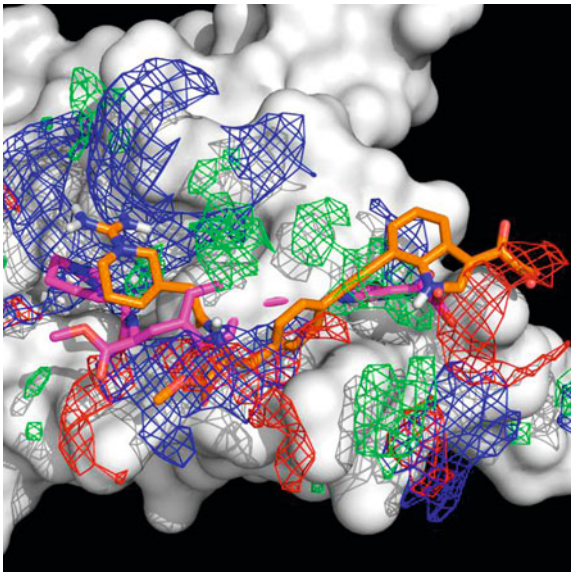


Fig. 3.5 The IL-2 receptor (*white surface*) and FLAP-GBPM interaction maps showing the hydrophobic (*green*), hydrogen bond donor (*blue*), and hydrogen bond acceptor (*red*) interacting regions as isocontours. After protein alignment, the ligand from 1M49 is shown in magenta; the surface from 1Z92 overlaps with the ligand in the central part of the figure, where induced fit has taken place in the 1M49 structure. The pose predicted by FLAP-GBPM is shown in orange

is therefore a limitation of the current approach, although the discrimination between actives and decoys was still excellent.

3.3 Concluding Remarks

The importance of protein–protein interactions as therapeutic targets is growing exponentially, and effective small molecule disruptors of these PPIs are being developed which will only reinforce this importance in the coming years. Methodologies aimed at supporting the discovery of such small molecule PPI inhibitors are also advancing; however, relevant in silico methods are still in their early years as researchers adapt methodologies from more traditional enzyme and GPCR inhibition approaches. Within this context, we have applied one of the earliest in silico methods, the GRID molecular interaction fields, to characterise protein–protein interactions using the GBPM approach together with the FLAP pharmacophoric field similarity approach. The combined FLAP-GBPM method allows the key protein–protein interacting regions to be identified without user bias, and small molecules are 'docked' and scored into the interaction site using the common MIF reference framework. The approach has shown great promise in several examples, although there are still areas to be explored; predicting the interacting regions without the structure of the interacting partner, and allowing receptor flexibility to take into account, the induced fit phenomena are two obvious examples that we are currently investigating.

Acknowledgments The authors would like to thank Jon Mason and Thierry Langer for their involvement in the original FLAP and GBPM methodology.

References

1. Toogood PL (2002) Inhibition of protein–protein association by small molecules: approaches and progress. J Med Chem 45:1543–1558
2. Wilson AJ (2009) Inhibition of protein–protein interactions using designed molecules. Chem Soc Rev 38:3289–3300
3. Wells JA, McLendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. Nature 450:1001–1009
4. Yin H, Hamilton AD (2005) Strategies for targeting Protein–Protein interactions with synthetic agents. Angew Chem Int Ed, 44, 4130–4163
5. Stumpf MPH, Thorne T, de Silva E, Stewart R, An HJ, Lappe M, Wiuf C (2008) Estimating the size of the human interactome. Proc Natl Acad Sci USA 105:6959–6964
6. Amaral LAN (2008) A truer measure of our ignorance. Proc Natl Acad Sci USA 105:6795–6796
7. Arkin MR, Wells JA (2004) Small molecule inhibitors of protein-protien interactions: progressing towards the dream. Nat Rev Drug Discovery 3:301–317
8. Hudis CA (2007) Trastuzumab–Mechanism of action and use in clinical practice. New Engl J Med 357:39–51

9. Sia SK, Carr PA, Cochran AG, Malashkevich VN, Kim PS (2002) Short constrained peptides that inhibit HIV-1 entry. Proc Natl Acad Sci USA 99:14664–14669
10. Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S-C, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, Rosenberg SH (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature 435:677–681
11. Vassilev LT, Vu BT, Graves B (2004) Carvaja, l D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., Fotouhi, N., Liu, E. A.: In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. Science 303:844–848
12. Higueruelo AP, Schreyer A, Bickerton GRJ, Pitt WR, Groom CR, Blundell TL (2009) Atomic interactions and profile of small molecules disrupting protein–protein interfaces: the TIMBAL Database. Chem Biol Drug Des 74:457–467
13. Robin WS (1998) High-throughput screening of historic collections: observations on file size, biological targets, and file diversity. Biotechnol Bioeng 61:61–67
14. Cochran AG (2000) Antagonists of protein–protein interactions. Chem Biol 7:R85–R94
15. Gandhi L, Camidge DR, de Oliveira MR, Bonomi P, Gandara D, Khaira D, Hann CL, McKeegan EM, Litvinovich E, Hemken PM, Dive C, Enschede SH, Nolan C, Chiu Y-L, Busman T, Xiong H, Krivoshik AP, Humerickhouse R, Shapiro GI, Rudin CM (2011) Phase I study of navitoclax (ABT-263), a Novel Bcl-2 family inhibitor, in patients with small-cell lung cancer and other solid tumors. J Clin Oncology 29:909–916
16. Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 28:849–857
17. Ortuso F, Langer T, Alcaro S (2006) GBPM: GRID-based pharmacophore model: concept and application studies to protein–protein recognition. Bioinformatics 22:1449–1455
18. Baroni M, Cruciani G, Sciabola S, Perruccio F, Mason JS (2007) A common reference framework for analyzing/comparing ligands and proteins. Fingerprints for ligands and proteins (FLAP): theory and application. J Chem Inf Model 47:279–294
19. Grigoriev A (2001) A relationship between gene expression and protein interactions on the proteome scale: Analysis of the bacteriophage T7 and the yeast Saccharomyces cerevisiae. Nucleic Acids Res 29:3513–3519
20. Ge H, Liu Z, Church GM, Vidal M (2001) Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. Nat Genet 29:482–486
21. Jansen R, Greenbaum D, Gerstein M (2002) Relating whole-genome expression data with protein–protein interactions. Genome Res 12:37–46
22. Skrabanek L, Saini HK, Bader GD, Enright AJ (2008) Computational prediction of protein–protein interactions. Mol Biotechnol 38:1–17
23. Gallet X, Charloteaux B, Thomas A, Brasseur R (2000) A fast method to predict protein interaction sites from sequences. J Mol Biol 302:917–926
24. Korn AP, Burnett RM (1991) Distribution and complementarity of hydropathy in multisubunit proteins. Proteins-Struct Funct Genetics 9:37–55
25. Young L, Jernigan RL, Covell DG (1994) A role for surface hydrophobicity in protein–protein recognition. Protein Sci 3:717–729
26. Mueller TD, Feigon J (2002) Solution structures of UBA domains reveal a conserved hydrophobic surface for protein–protein interactions. J Mol Biol 319:1243–1255
27. Ofran Y, Rost B (2003) Analysing six types of protein–protein interfaces. J Mol Biol 325:377–387
28. Jones S, Thornton JM (1997) Prediction of protein–protein interaction sites using patch analysis. J Mol Biol 272:133–143
29. Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJ, Vajda S, Vakser I, Wodak SJ (2003) CAPRI: a critical assessment of predicted interactions. Proteins 52:2–9
30. Zhou HX, Shan YB (2001) Prediction of protein interaction sites from sequence profile and residue neighbor list. Proteins-Struct Funct Genet 44:336–343

31. Fariselli P, Pazos F, Valencia A, Casadio R (2002) Prediction of protein–protein interaction sites in heterocomplexes with neural networks. Eur J Biochem 269:1356–1361
32. Ofran Y, Rost B (2003) Predicted protein–protein interaction sites from local sequence information. FEBS Lett 544:236–239
33. De Genst E, Areskoug D, Decanniere K, Muyldermans S, Andersson K (2002) Kinetic and affinity predictions of a protein–protein interaction using multivariate experimental design. J Biol Chem 277:29897–29907
34. Thorn K, Bogan A (2001) ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. Bioinformatics 3:284–285
35. Fischer T, Arunachalam K, Bailey D, Mangual V, Bakhru S, Russo R, Huang D, Paczkowski M, Lalchandani V, Ramachandra C (2003) The binding interface database (BID): a compilation of amino acid hot spots in protein interfaces. Bioinformatics 11:1453–1454
36. Bogan A, Thorn K (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280:1–9
37. Moreira I, Fernandes P, Ramos M (2007) Hot spots-A review of the protein–protein interface determinant amino-acid residues. Proteins 68:803–812
38. Li J, Liu Q (2009) 'Double water exclusion': a hypothesis refining the O-ring theory for the hot spots at protein interfaces. Bioinformatics 25:743–750
39. Kruger DM, Garzon JI, Montes PC, Gohlke H (2011) Predicting protein–protein interactions with DrugScorePPI: fully flexible docking, scoring, and in silico alanine scanning. J. Cheminf 3:P36
40. Kruger DM, Gohlke H (2010) DrugScorePPI webserver: fast and accurate in silico alanine scanning for scoring protein–protein interactions. Nucleic Acids Res 38:W480–W486
41. Lise S, Archambeau C, Pontil M, Jones DT (2009) Prediction of hot spot residues at protein–protein interfaces by combining machine learning and energy-based methods. BMC Bioinformatics 10:365–382
42. Lise S, Buchan D, Pontil M, Jones DT (2011) Predictions of hot spot residues at Protein–Protein interfaces using support vector machines. PLoS ONE 6:e16774
43. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) The FoldX web server: an online force field. Nucleic Acids Res 33:W382–W388
44. Ofran Y, Rost B (2007) Protein-protein interaction hotspots carved into sequences. PLoS Comput Biol 3:e119
45. Darnell S, Page D, Mitchell J (2007) An automated decision-tree approach to predicting protein interaction hot spots. Proteins 68:813–823
46. Darnell S, LeGault L, Mitchell J (2008) KFC server: interactive forecasting of protein interaction hot spots. Nucleic Acids Research, W265–W269
47. Guney E, Tuncbag N, Keskin O, Gursoy A (2008) HotSprint: database of computational hot spots in protein interfaces. Nucleic Acids Research, D662–D666
48. Tuncbag N, Gursoy A, Keskin O (2009) Identification of computational hot spots in protein interfaces: combining solvent accessibility and interresidue potentials improves the accuracy. Bioinformatics 1513–1520
49. Cho K, Kim D, Lee D (2009) A feature-based approach to modeling proteinprotein interaction hot spots. Nucleic Acids Res 37:2672–2687
50. Xia J-F, Zhao X-M, Song J, Huang D (2010) APIS: acurate prediction of hot spots in protein interfaces by combining protrusion index with solvent accessibility. BMC Bioinfo 11:174–188
51. Landon MR, Lancia DR Jr, Yu J, Thiel SC, Vajda S (2007) Identification of hot spots within druggable binding sites of proteins by computational solvent mapping. J Med Chem 50:1231–1240
52. Brenke R, Kozakov D, Chuang G-Y, Beglov D, Hall D, Landon MR, Mattos C, Vajda S (2009) Fragment-based identification of druggable "hot spots" of proteins using Fourier domain correlation techniques. Bioinformatics 25:621–627
53. Kozakov D, Hall DR, Chuang G-Y, Cencic R, Brenke R, Grove LE, Beglov D, Pelletier J, Whitty A, Vajda S (2011) Structural conservation of druggable hot spots in protein–protein interfaces. Proc Natl Acad Sci USA 108:13528–13533

54. Dennis S, Kortvelyesi T, Vajda S (2002) Computational mapping identifies the binding sites of organic solvents on proteins. Proc Natl Acad Sci USA 99:4290–4295
55. www.moldiscovery.com
56. Goodford P (2006) The basic principles of GRID, in molecular interaction fields, Cruciani G (ed) Wiley, pp 3–26
57. Wade RC (2006) Calculation and application of molecular interaction fields, in molecular interaction fields, Cruciani G (ed) Wiley, pp 27–42
58. Cross S, Cruciani G (2010) Molecular fields in drug discovery: getting old or reaching maturity? Drug Discovery Today 15:23–32
59. Von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dyason JC, Jin B, Phan TV, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Woods JM, Bethel RC, Hotham VJ, Cameron JM, Penn CR (1993) Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 363:418–423
60. Fox T (2006) Protein selectivity studies using GRID-MIFs, in molecular interaction fields, Cruciani G (ed), Wiley, pp 45–82
61. Cruciani G, Pastor M, Guba W (2000) VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. Eur J Pharm Sci 11:S29–S39
62. Cruciani G, Carosati E, De Boeck B, Ethirajulu K, Mackie C, Howe T, Vianello R (2005) MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. J Med Chem 48:6970–6979
63. Milletti F, Storchi L, Sforna G, Cruciani G (2007) New and original pKa prediction method using grid molecular interaction fields. J Chem Inf Model 47:2172–2181
64. Bergmann R, Linusson A, Zamora I (2007) SHOP: scaffold HOPping by GRID-based similarity searches. J Med Chem 50:2708–2717
65. Carosati E, Cruciani G, Chiarini A, Budriesi R, Ioan P, Spisani R, Spinelli D, Cosimelli B, Fusi F, Frosini M, Matucci R, Gasparini F, Ciogli A, Stephens PJ, Devlin FJ (2006) Calcium channel antagonists discovered by a multidisciplinary approach. J Med Chem 49:5206–5216
66. Huang N, Shoichet BK, Irwin J (2006) Benchmarking sets for molecular docking. J Med Chem 49:6789–6801
67. Cross S, Baroni M, Carosati E, Benedetti P, Clementi S (2010) FLAP: GRID molecular interaction fields in virtual screening validation using the DUD data set. J Chem Inf Model 50:1442–1450
68. Patel Y, Gillet VJ, Bravi G, Leach AR A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. J Comp-Aided Mol Des, 16, 653–681
69. Cross S, Cruciani G unpublished results
70. Alcaro S, Artese A, Ceccherini-Silberstein F, Chiarella V, Dimonte S, Ortuso F, Perno CF (2010) Computational analysis of human immunodeficiency virus (HIV) Type-1 reverse transcriptase crystallographic models based on significant conserved residues found in highly active antiretroviral therapy (HAART)-treated patients. Curr Med Chem 17:290–308
71. Accelrys 2003 http://www.accelrys.com
72. Wu G, Chai J, Suber TL, Wu JW, Du C, Wang X, Shi Y (2000) Structural basis of IAP recognition by Smac/DIABLO. Nature 408:1008–1012
73. Shiozaki EN, Chai J, Rigotti DJ, Riedl SJ, Li P, Srinivasula SM, Alnemri ES, Fairman R, Shi Y (2003) Mechanism of XIAP-mediated inhibition of caspase-9. Mol Cell 11:519–527
74. Oost TK, Sun Ch, Armstrong RC, Al-Assaad A-S, Betz SF, Decworth TL, Ding H, Elmore SW, Meadows RP, Olejniczak ET, Oleksijew A, Oltersdorf T, Rosenberrg SH, Shoemaker AR, Tomaselli KJ, Zou H, Fesik SW (2004) Discovery of potent antagonists of the antiapoptotic protein XIAP for the treatment of cancer. J Med Chem 47:4417–4426
75. Liu Z, Sun C, Olejniczak ET, Meadows RP, Betz SF, Oost T, Herrmann J, Wu JC, Fesik SW (2000) Structural basis for binding of Smac/DIABLO to the XIAP BIR3 domain. Nature 408:1004–1008
76. Sun H, Stuckey JA, Nikolovska-Coleska Z, Qin D, Meagher JL, Qiu S, Lu J, Yang C, Saito NG, Wang S (2008) Structure-based design, synthesis, evaluation, and crystallographic

studies of conformationally constrained Smac mimetics as inhibitors of the X-linked inhibitor of apoptosis protein (XIAP). J Med Chem 51:7169–7180
77. Wist AD, Gu L, Riedl SJ, Shi Y, McLendon GL (2007) Structure-activity based study of the Smac-binding pocket within the BIR3 domain of XIAP. Bioorg Med Chem 15:2935–2943
78. Nikolovska-Coleska Z, Meagher JL, Jiang S, Yang CY, Qiu S, Roller PP, Stuckey JA, Wang S (2008) Interaction of a cyclic, bivalent smac mimetic with the x-linked inhibitor of apoptosis protein. Biochemistry 47:9811–9824
79. Cossu F, Mastrangelo E, Milani M, Sorrentino G, Lecis D, Delia D, Manzoni L, Seneci P, Scolastico C, Bolognesi M (2009) Designing smac-mimetics as antagonists of XIAP, cIAP1, and cIAP2. Biochem Biophys Res Commun 378:162–167
80. Mastrangelo E, Cossu F, Milani M, Sorrentino G, Lecis D, Delia D, Manzoni L, Drago C, Seneci P, Scolastico C, Rizzo V, Bolognesi M (2008) Targeting the X-linked inhibitor of apoptosis protein through 4-substituted azabicyclo [5.3.0]alkane smac mimetics. Structure, activity, and recognition principles. J Mol Biol 384:673–689
81. Cossu F, Milani M, Mastrangelo E, Vachette P, Servida F, Lecis D, Canevari G, Delia D, Drago C, Rizzo V, Manzoni L, Seneci P, Scolastico C, Bolognesi M (2009) Structural basis for bivalent Smac-mimetics recognition in the IAP protein family. J Mol Biol 392:630–644
82. Ndubaku C, Varfolomeev E, Wang L, Zobel K, Lau K, Elliott LO, Maurer B, Fedorova AV, Dynek JN, Koehler M, Hymowitz SG, Tsui V, Deshayes K, Fairbrother WJ, Flygare JA, Vucic D (2009) Antagonism of c-IAP and XIAP proteins is required for efficient induction of cell death by small-molecule IAP antagonists. ACS Chem Biol 4:557–566
83. Levine AJ, Hu W, Feng Z (2006) The p53 pathway: what questions remain to be explored? Cell Death Differ 13:1027–1036
84. Grasberger BL, Lu T, Schubert C, Parks DJ, Carver TE, Koblish HK, Cummings MD, LaFrance LV, Milkiewicz KL (2005) Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. J Med Chem 48:909–912
85. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303:844–848
86. Fry DC, Emerson SD, Palme S, Vu BT, Liu CM, Podlaski F (2004) NMR structure of a complex between MDM2 and a small molecule inhibitor. J Biomol NMR 30:163–173
87. Fasan R, Dias RL, Moehle K, Zerbe O, Obrecht D, Mittl PR, Robinson JA (2006) Structure-activity studies in a family of beta-hairpin protein epitope mimetic inhibitors of the p53-HDM2 protein–protein interaction. ChemBioChem 7:515–526
88. Sakurai K, Schubert C, Kahne D (2006) Crystallographic analysis of an 8-mer p53 peptide analogue complexed with MDM2. J Am Chem Soc 128:11000–11001
89. Pazgier M, Liu M, Zou G, Yuan W, Li C, Li C, Li J, Monbo J, Zella D, Tarasov SG, Lu W (2009) Structural basis for high-affinity peptide inhibition of p53 interactions with MDM2 and MDMX. Proc Natl Acad Sci USA 106:4665–4670
90. Czarna A, Popowicz GM, Pecak A, Wolf S, Dubin G, Holak TA (2009) Hot, hotter, hottest. Cell Cycle 8:1176–1184
91. Li C, Pazgier M, Liu M, Lu WY, Lu W (2009) Apamin as a template for structure-based rational design of potent peptide activators of p53. Angew.Chem.Int.Ed.Engl 48, 8712–8715
92. Liu M, Li C, Pazgier M, Li C, Mao Y, Lv Y, Gu B, Wei G, Yuan W, Zhan C, Lu WY, Lu W (2010) D-peptide inhibitors of the p53-MDM2 interaction for targeted molecular therapy of malignant neoplasms. AngewProc.Natl.Acad.Sci.USA, 398, 200–213
93. Popowicz GM, Czarna A, Wolf S, Wang K, Wang W, Domling A, Holak TA (2010) Structures of low molecular weight inhibitors bound to MDMX and MDM2 reveal new approaches for p53-MDMX/MDM2 antagonist drug discovery. Cell Cycle 9:1104–1111
94. Liu M, Pazgier M, Li C, Yuan W, Li C, Lu W A (2010) left-handed solution to peptide inhibition of the p53-MDM2 interaction. Angew Chem Int Ed Engl 49, 3649–3652
95. Nelson BH, Willerford DM (1998) Biology of the Interleukin-2 Receptor. Adv Immunol 70:1–81

96. Rickert M, Wang X, Boulanger MJ, Goriatcheva N, Garcia KC (2005) The Structure of Interleukin-2 Complexed with Its Alpha Receptor. Science 308:1477–1480
97. Arkin MA, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, Raphael DR, Taylor L, Wang J, McDowell RS, Wells JA, Braisted AC (2003) Binding of small molecules to an adaptive protein–protein interface. Proc Natl Acad Sci USA 100:1603–1608
98. Thanos CD, Randal M, Wells JA (2003) Potent small-molecule binding to a dynamic hot spot on IL-2. J Am Chem Soc 125:15280–15281

Chapter 4
NMR as a Tool to Target Protein–Protein Interactions

Rebecca Del Conte, Daniela Lalli and Paola Turano

4.1 Introduction

NMR spectroscopy plays a dual role in projects aimed at targeting protein–protein interactions (PPIs).While it has been extensively validated as an efficient technique for the initial screening and identification of weakly interacting fragments and for subsequently guiding their optimization into molecules with higher affinity and more favorable drug-like properties, it also represents an extremely powerful tool to monitor the formation of protein–protein complexes in solution and to obtain structural information on these adducts. It allows the identification of the protein interfaces and, in some cases, provides intermolecular distances and orientational restraints that lead to the definition of the relative arrangement of the two proteins. In particular, it constitutes the structural technique of choice for studying weak/transient protein–protein interactions, which represent the natural targets for drug discovery projects addressing PPIs.

4.2 Thermodynamics and Kinetics of PPIs

The "interactome" includes the totality of the PPIs in a living cell. These interactions span an affinity range that is extremely broad, with dissociation constants K_d from 10^{-2} to 10^{-16} M. In a recent review [1], such interactions have been

R. Del Conte · D. Lalli · P. Turano (✉)
Magnetic Resonance Center (CERM), University of Florence,
Via Luigi Sacconi 6, 50019 Florence, Sesto Fiorentino, Italy
e-mail: turano@cerm.unifi.it

R. Del Conte
e-mail: delconte@cerm.unifi.it

D. Lalli
e-mail: lalli@cerm.unifi.it

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4_4,

categorized into strong (with dissociation constant $K_d < 1$ μM), weak (1 μM < $K_d < 100$ μM) and ultra weak ($K_d > 100$ μM). In the assumption of a one-step reaction, K_d equals the ratio k_{off}/k_{on}, where k_{off} and k_{on} are the dissociation and association rate constants, respectively. Different combinations of k_{off} and k_{on} can give rise to similar affinities. A wide spectrum of association rate constants has been reported. The rate of association of a protein complex is limited by diffusion and geometric constraints of the binding sites: in order two proteins recognize each other, they have to be oriented with high specificity [2]. The presence of orientational constraints for the formation of complexes that do not involve a significant electrostatic contribution corresponds to k_{on} values in the range $\sim 10^5$–10^6 M^{-1} s^{-1}. Higher association rates implicate electrostatic enhancement resulting from complementary electrostatic surfaces. Electrostatic interactions are long-range acting and therefore accelerate the recognition process beyond a merely diffusion-limited situation, with k_{on} values up to 10^8–10^9 M^{-1} s^{-1}. Diffusion-controlled or electrostatic-guided associations typically involve only local conformational changes upon complex formation. In some cases, however, the establishment of an adduct involves gross conformational changes such as loop reorganization or domain movement, which slow the association rate constant down to 10^3 M^{-1} s^{-1} [2, 3]. Typical k_{off} is in the 10^4–10^{-7} s^{-1} range [4, 5]. k_{off} is directly related to the lifetime of the complex: the mean lifetime is given by $1/k_{off}$, whereas the half-life of the complex is $\ln 2/k_{off}$ (i.e., $0.69315/k_{off}$) [6]. A higher value of k_{off} means a shorter lifetime for the complex and vice versa.

Transient complexes form when a high turnover is a functional requirement and their components associate and dissociate rapidly, namely with $k_{off} \geq 10^3$ s^{-1} and k_{on} in the range of 10^7–10^9 M^{-1} s^{-1}. This results in dissociation constants typical for weak and ultra-weak complexes and lifetimes $\leq$ ms. Revealing the presence of such interactions is experimentally challenging because they do not result in a sufficient amount of complexes that can be directly measured. NMR spectroscopy has the ability to detect and provide structural information on weak/ultra-weak interactions on transient PPIs, thus leading to the identification of the interaction area to be targeted in structure-based discovery of new small molecule inhibitors [7–11].

4.3 NMR for the Characterization of Weak/Ultra-Weak PPIs

The parameter relevant for the interpretation of the binding dynamics in NMR experiments is the exchange rate constant, k_{ex}. It depends upon the relative populations of the protein and ligand, as well as the two exchange rates k_{on} and k_{off} (and therefore upon the dissociation constant K_d), as expressed by the relationship $k_{ex} = k_{on}[P] + k_{off}$, where $[P]$ is the concentration of the free protein. The value of the exchange rate constant defines the different situations that can be encountered in NMR binding experiments. The k_{ex} value has to be compared with the chemical shift difference $\Delta\nu$ (expressed in unit of frequency, i.e., Hz $= s^{-1}$) of signals in the

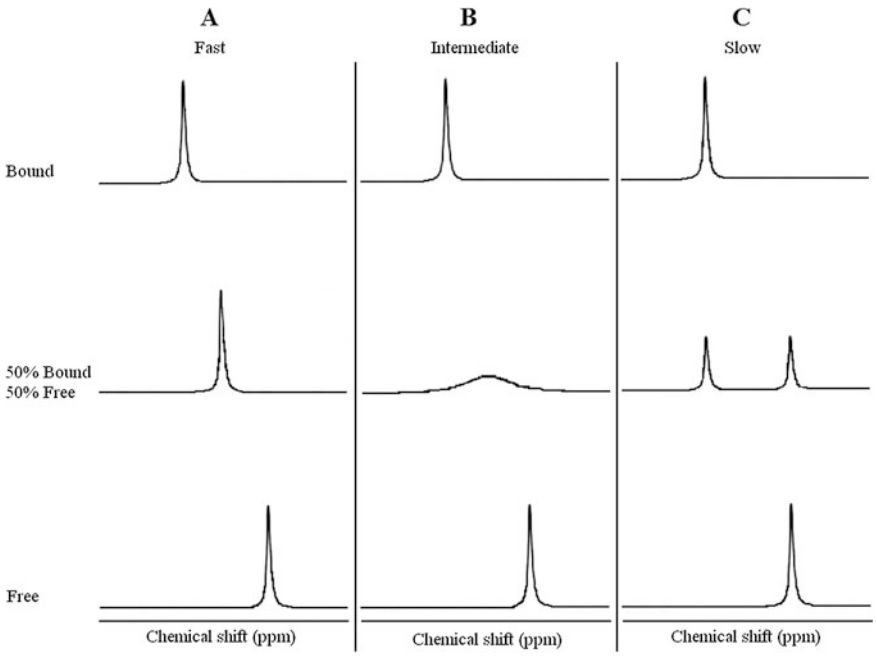


Fig. 4.1 Schematic representation of the monodimensional NMR spectrum for a nucleus experiencing two distinct chemical environments in the free and bound form of the molecule to which it belongs. Two limit situations are possible: the fast exchange regime (*panel A*) and the slow exchange regime (*panel C*). When k_{ex} is much faster than the difference in chemical shift (expressed in frequency units), only one signal is observed at a population-weighted average chemical shift (*A*). When k_{ex} is much slower than the difference in chemical shift, two separate peaks are observed at the frequency of the free and bound forms, with relative intensities corresponding to the relative populations (*C*). When k_{ex} is of the order of the difference in chemical shift, we are under intermediate exchange regimes and the linewidths are increased via exchange broadening (*B*)

free and bound state (Fig. 4.1). If $k_{ex} \gg |\Delta v|$, the system is under the fast exchange regime on the chemical shift time scale, and the experimental chemical shift value for a given nucleus will correspond to the population-weighted average of its chemical shift in the bound and free state of the molecule (Fig. 4.1A). In the fast exchange regime, not only the chemical shift, but also all the other NMR parameters result population weighted. Slow exchange on the chemical shift time scale occurs when $k_{ex} \ll |\Delta v|$; two separate signals are observed for the same nucleus, one for the bound and one for the free state, with relative intensities that correspond to the relative concentrations of the two states (Fig. 4.1C). Intermediate exchange regimes give rise to signal broadening. For $k_{ex} \sim \Delta v$, only one signal is observed with the population-weighted average chemical shift; its linewidth is increased because of the phenomenon called "exchange broadening" and such an effect can be so severe to broaden the signal beyond detection (Fig. 4.1B). Very roughly, in the diffusion-controlled association regime, dissociation constants

(K_d) in the millimolar range correspond to fast exchange, in the micromolar range to intermediate exchange and in the nanomolar range to slow exchange regime[12, 13].

4.3.1 Chemical Shift Perturbation

Changes in chemical shift between the free and bound state of a molecule reflect the change in the chemical environment of the corresponding nucleus. In the absence of conformational rearrangements, they are confined to nuclei of residues in the interaction area. In protein interaction studies, the $^{1}H–^{15}N$ heteronuclear single-quantum correlation (HSQC) experiment represents the gold standard for monitoring intermolecular interactions (Fig. 4.2a). Backbone amide protons are the most sensitive probes to variations in the chemical environment due to the intermolecular interaction and are therefore able to give rise to measurable effects even in the case of ultra-weak interactions (up to mM K_d's) [14, 15]. Moreover, the $^{1}H–^{15}N$ HSQC is a simple and fast experiment that with the cryoprobe technology can be acquired in a few minutes even for very low protein concentrations (down to tens of μM). In interaction experiments, the $^{1}H–^{15}N$ HSQC spectrum of one protein is monitored when the unlabeled interaction partner is titrated in, and the chemical shift perturbations are recorded for each amino acid (Fig. 4.2b). Provided the assignment of the $^{1}H–^{15}N$ HSQC spectrum is available, shift perturbation measurements identify residues at the interface [16] (Fig. 4.2c). If a structural model of the protein exists, residues undergoing meaningful chemical shift perturbations are mapped on the protein structure, thus providing the location of the interface (Fig. 4.2c). In the case of protein–protein interactions, the procedure is repeated for the second partner. Chemical shift perturbation mapping yields the interaction area on the individual binding partners, but does not allow defining the relative orientation of the two molecules nor the atom-to-atom interactions at the basis of the recognition process. Nevertheless, residues undergoing chemical shift perturbations can be used as selection filters in data-driven computational soft-docking programs [17, 18].

In the presence of binding-induced conformational rearrangements, the chemical shift perturbation may extend to residues that are not at the interface, and the chemical shift perturbation fails as a mapping tool, although it still represents an excellent indicator of allosteric processes. Such a situation is associated with slow k_{on}.

Finally, under fast exchange conditions, K_d values and stoichiometry can be estimated from the titration of the ^{15}N-labeled observed protein with its unlabeled binding partner. Resonances of the nuclei at the interface shift in a continuous fashion in the HSQC spectra during the titration and fitting the chemical shift perturbation as a function of the concentration of the partner provide affinity constants (Fig. 4.3). If chemical shift changes in the HSQC maps are linear and occur at the same rate for the affected residues, a single binding event is occurring;

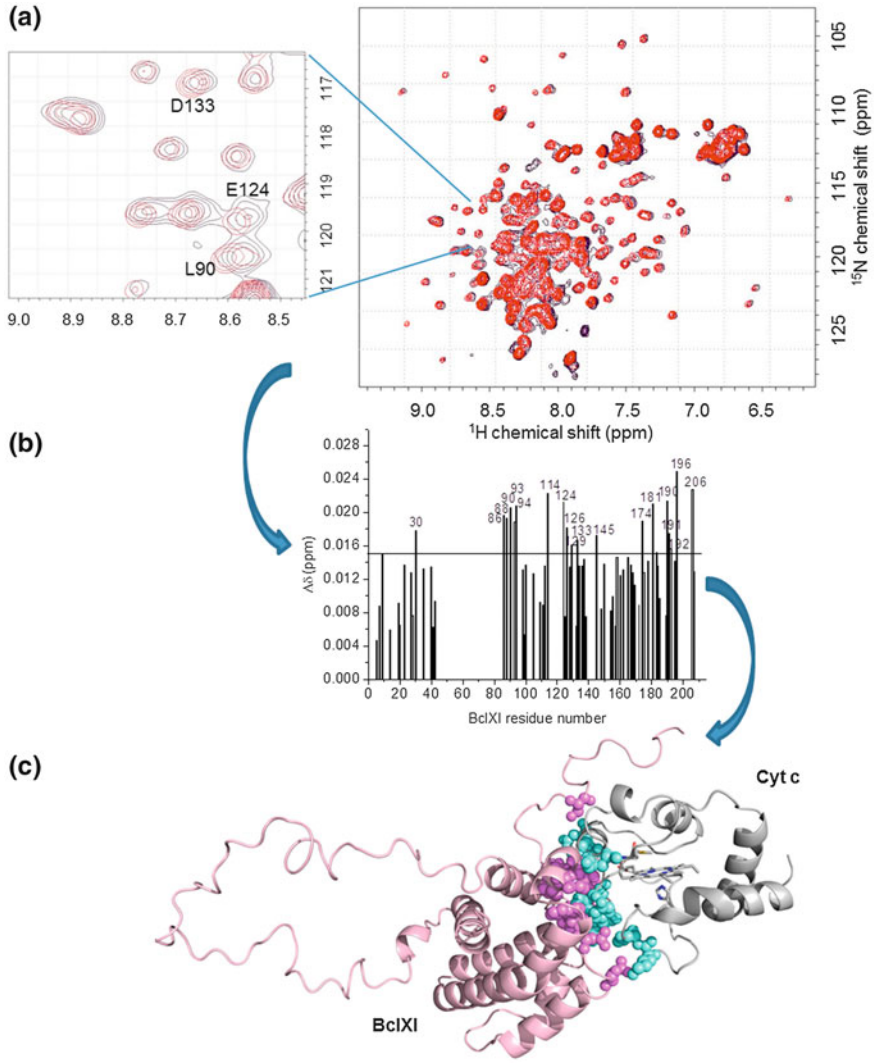


Fig. 4.2 Workflow for PPI characterization by chemical shift mapping using as case example the Bcl-Xl-cytochrome c complex [97]. **a** Chemical shift perturbations in the ^{1}H–^{15}N HSQC spectrum of ^{15}N-enriched Bcl-Xl upon titration with unlabeled human cytochrome c are observed for a number of residues. The inset reports an enlargement of the spectral region containing the resonances of residues L90, E124 and D133. **b** The Garrett plot reports the weighted average chemical shift change for ^{1}H and ^{15}N between the initial and the final stage of a titration as a function of the residue number. Meaningful chemical shift perturbations are those above a selected threshold, here provided by the *horizontal line*. **c** Mapping of the chemical shift perturbations on the structure of Bcl-Xl-cytochrome c: residues above the threshold in the Garrett plot are indicated by *magenta* spheres. An analogous procedure, based on the titration of ^{15}N-enriched human cytochrome c with unlabeled Bcl-Xl, provides the chemical shift mapping of the interaction area (*cyan sphere*). Data-driven docking procedures provided the structural model of the complex here reported. Adapted from Ref. [97]

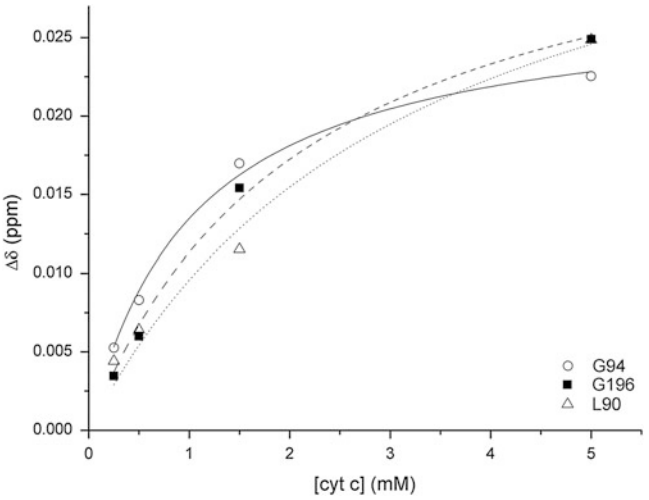


Fig. 4.3 A plot of the weighted average chemical shift changes of the backbone amide resonances for selected HSQC peaks (L90, G94, G196) in ^{15}N-enriched Bcl-Xl as a function of the concentration of human cytochrome c. Fitting of these data provides the K_d value

if changes for different residues occur at different rates and/or trajectories deviate from linearity, more than one binding event/site is implicated. Under the slow exchange regime, during a titration, the set of resonances associated with the free state will decrease in intensity, while that of the bound state will increase in intensity. For residues that are not affected by the interaction, the two sets will coincide. Disappearance of a resonance of the free state from its original position indicates that the corresponding residue is affected by the interaction. An evaluation of the extent of the chemical shift change upon binding will require an independent assignment procedure for the bound state. The affinity constant can be quantitated by fitting the intensity of the disappearing or appearing peaks as a function of the concentration of the added titrant. Intermediate exchange regime is accompanied by resonance broadening, sometimes beyond detection. While observation of this phenomenon provides a proof of the binding event, frequencies of the affected resonances become poorly defined and a detailed spectral analysis may become unfeasible.

4.3.2 Changes in Protein Dynamics

Interaction between two molecules not only implies changes in chemical environment due to local/extended conformational variations, but also affects the solvent accessibility and internal protein mobility.

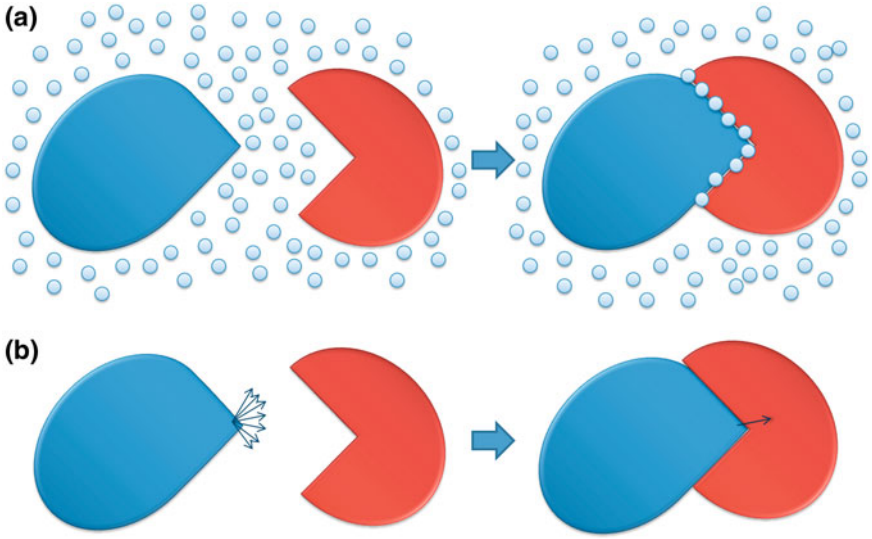


Fig. 4.4 Cartoon representation of the interaction between two proteins (*blue* and *red* objects). **a** Complex formation is always accompanied by reduced solvent accessibility of the contact area with exclusion of bulk water. Some solvent molecules may remain at the interface: they can be bridging, that is, having significant interactions with both proteins; nonbridging, that is, having significant interactions with only one of the two proteins; or simply trapped without significant interactions with either protein. **b** The side chain (exemplified by an *arrow*) of a residue of the *blue* protein in its free form exists in a number of energetically similar conformations that can be represented by a dynamic ensemble of interconverting states. Upon binding, the side chain assumes a single conformation that often corresponds to one of those of the ensemble of the free form (conformational selection), but in some cases may correspond to a new conformation that would not spontaneously adopt in its unligated state. Either situation causes a reduction in side chain flexibility

Often, establishment of intermolecular contacts shields the interaction area from exchange with bulk solvent (Fig. 4.4a), and therefore, mapping with amide proton exchange (i.e., by measuring the solvent accessibility of each amide group in the free and bound state) can further contribute to the identification of residues at the surface area involved in the recognition process.

In most cases, internal dynamics become quenched at the interface: regions prone to give rise to interactions in the presence of partners are often intrinsically dynamic in the free state and inclined to induced fit conformational changes. Upon interaction, one conformation or subset of conformations becomes stabilized, thus changing the population of the possible states with respect to the free form [13] (Fig. 4.4b).

4.3.3 Changes in the Overall Tumbling Time

Establishment of intermolecular interactions gives rise to changes in the size of the system under study.

For protein–protein interactions, an overall increase in transverse relaxation rate values is observed, which is consistent with an increase in the overall tumbling correlation time upon complex formation. In the fast exchange regime, the measured transverse relaxation rate is an average of the transverse relaxation rates of the free and bound protein forms, weighted by their molar fraction. This phenomenon can be used to assess complex formation, but also causes signal broadening in the spectra of the bound form of a protein, thus often requiring *ad hoc* experimental procedures for the obtainment of $^{1}H–^{15}N$ correlations at the basis of the chemical shift perturbation mapping approach, as detailed below.

Interaction of a protein with another protein greatly increases the size of the system to be characterized. With increasing molecular mass, the slow molecular tumbling of the complex in solution causes fast transverse relaxation that broadens signals beyond detection in standard $^{1}H–^{15}N$ HSQC experiments. For total MW < 80 kDa, combination of transverse relaxation optimized spectroscopy (TROSY) [19] and isotope enrichment with ^{2}H for nonexchangeable hydrogens alleviates line broadening effects. For supramolecular assemblies larger than this threshold, detection of $^{1}H–^{15}N$ correlations requires the use of cross-relaxation-enhanced polarization transfer (CRINEPT)-TROSY or cross-correlated relaxation-induced polarization transfer (CRIPT)-TROSY experiments [20]. Such high total molecular weight can be accomplished by combining proteins of different sizes. If one partner molecule has a MW >80–100 kDa by itself, practical applications are limited to correlation experiments for chemical shift measurement only of the low molecular weight component. Under these circumstances, when a relatively low molecular weight component interacts with a larger system, the signals of the former (in a ^{15}N, >90 $\%^{2}H$-labeled form) are detectable although much broader than in its free form; this approach has allowed "monolateral" chemical shift mapping for complexes up to 1 MDa [21-23], taking advantage of the assignment obtained for the free state. In the spectrum of the large molecular weight component, signal overlap and signal broadening are too severe and no suitable NMR experiments exist at the moment that may allow assignment of backbone amides even in the free state, thus making the chemical shift mapping impossible [21, 24].

4.3.4 Distance and Orientational Restraints

The information on the binding interfaces obtainable as described above, might be complemented with data restraining the distance and relative orientation of the individual proteins using (1) nuclear Overhauser effect (NOE), (2) residual dipolar couplings (RDCs), (3) paramagnetic relaxation enhancement (PRE) and pseudo-contact shifts (PCSs).

(1) *Nuclear Overhauser Effect*. The gold standard for the obtainment of structural information in protein NMR is the Nuclear Overhauser Effect SpectroscopY (NOESY) experiment, in either its 2D homonuclear version or the ^{15}N- or ^{13}C edited 3D variants, which provides ^{1}H–^{1}H cross-peaks for all pairs of protons that fall at short distance (within about 5 Å) one from the other. In a rigid system, translation of NOE intensities into distances is possible thanks to the linear correlation existing between peak volume and $1/r^6$ (where r is the proton–proton distance). The extent of the NOE is also proportional to the correlation time for the dipolar interaction between the two nuclei. Any molecule in solution rotates as a whole, with a tumbling rate that depends on its size and shape as well as on the viscosity of the solution. Under these circumstances, the correlation time for the tumbling can reasonably be derived from the Stokes–Einstein–Debye equation [25] and lies in the range of several to hundreds of nanoseconds. Dipolar interactions across a binding interface can, in principle, give rise to NOEs, which can thus provide intermolecular distance restraints. However, it is often difficult to determine such NOEs. In the case of a complex, the correlation time for the dipolar interaction can be dominated by the exchange time between the free and bound forms. Dissociation rates faster than the tumbling are effective in reducing the dipolar interaction even down to values where the NOE effect is no more measurable. In practice, intermolecular NOEs to define the three-dimensional structure of a protein–protein complex are essentially restricted to tight complexes ($K_d < 10$ μM) [10]. Under these conditions, isotope edited and filtered NMR pulse sequences are used to distinguish between inter- and intramolecular NOEs [26, 27].

(2) *Residual Dipolar Couplings.* RDCs provide long-range structural restraints for NMR structure determination of macromolecules that are not accessible by most other NMR observables, which are dependent on close spatial proximity of nuclei (namely presence of chemical bonds giving rise to strong J-couplings or interproton distances within 5 Å providing measurable NOE effects). Proteins usually have nonzero, although small, magnetic susceptibility anisotropy and therefore can partially align at very high magnetic fields; the effect can be enhanced through addition in solution of alignment media such as bicelles or bacteriophages [28]. As a consequence of this partial alignment, dipolar interactions are not fully averaged out as it would be the case for a fully isotropic situation. In NMR spectra, RDCs appear as an additional contribution to the scalar *J* coupling splitting; their magnitude and sign must be determined by taking the difference in *J* values under anisotropic conditions and under isotropic conditions, or by taking the difference in *J* values in different alignment media. The measured effect is small, and specific experiments have been developed for the measurement of H–N or H–C RDCs. RDCs are effective in defining the relative orientation of pairs of nuclei (most commonly ^{1}H–^{15}N of backbone amides) within a molecular frame. RDCs can also be used for the purpose of deriving the relative orientation of two proteins

in a complex. This can be achieved by superimposing their alignment tensors independently determined. The structure of the complex can be obtained by keeping the structures of the individual proteins rigid. When the structures of the individual proteins/domains are known, the relative orientations of all dipolar vectors within the individual proteins are given. The measured dipolar couplings, in turn, reflect the orientation of the dipoles with respect to the magnetic field. It is then possible to carry out a computer procedure rotating these individual proteins to fit the measured dipolar couplings. Under fast exchange conditions, in order to overcome the difficulties encountered in obtaining RDCs that emanate from the complex alone, a titration approach can be employed where RDCs are measured in different equilibrium mixtures of the free and bound form. While the RDCs of the free states can be measured directly, the RDCs originating from the bound state will be obtained indirectly by extrapolation of the RDCs in the different equilibrium mixtures.

(3) *Paramagnetic derived constraints.* A number of metalloproteins contain paramagnetic metal ions or their native diamagnetic metal ion can be substituted with a paramagnetic one [29]; many other proteins can be functionalized adding paramagnetic tags [30]. The interaction between the unpaired electron(s) at the metal center and the surrounding nuclei provides further long-range structural information. Unpaired electrons and nuclei interact through two main mechanisms: a through-space dipolar interaction and a through-bond contact interaction; both of them affect the nuclear chemical shift and relaxation times. The whole effect on the chemical shift is called hyperfine shift; it can be positive or negative and adds to the diamagnetic contribution. It can be factorized in a contact shift and in a pseudocontact shift. The former is proportional to the unpaired electron spin density on the resonating nucleus and is effective only on the nuclei of the paramagnetic metal ion(s) ligands. The pseudocontact contribution to the hyperfine shift derives from a through-space interaction effective on all the nuclei within a certain distance from the paramagnetic metal ion(s). It originates from the dipolar coupling of the unpaired electron spin with the nuclear spin of surrounding atoms. If the magnetic moment of the electron is anisotropic (i.e., it has nonzero magnetic susceptibility anisotropy), the magnetic moment of the electron changes with orientation, the electron–nuclear dipolar coupling does not average to zero with molecular tumbling in solution and yields a shift that is dependent on both the electron–nuclear distance and the orientation of the electron–nucleus vector with respect to the magnetic susceptibility tensor. The pseudo contact shift is proportional to r^{-3}, where r is the electron–nucleus distance, in the approximation that the unpaired electron spin density is all metal centered, and therefore, the effect propagates to atoms far from the paramagnetic center. It also depends on angular parameters and thus provides information on the spatial position of the nucleus with respect to the magnetic axes frame that has the metal ion in its origin.

As far as relaxation times are concerned, paramagnetic contributions to relaxation add to the diamagnetic contributions. Paramagnetic systems have shorter relaxation time values (T_1 and T_2) with respect to their diamagnetic analogues because proximity of unpaired electron spin to a nuclear spin provides to nuclear spins further relaxation mechanisms. The paramagnetic contributions to the nuclear T_1 and T_2 relaxation times arise from several different terms: (1) electron–nuclear dipolar, which depends upon r^{-6}, where r is again the nucleus–metal ions distance in a metal-centered approximation; (2) contact, due to delocalized electron spin density and confined to metal ligands; (3) Curie spin, which arises from the interaction between the nuclear spin and the large time-averaged magnetic moment of the electron, again depends upon r^{-6}, and it becomes very important for high molecular weight systems at high magnetic fields. These paramagnetic effects can shorten nuclear T_1 and T_2 relaxation times by many orders of magnitude, and their extent depends upon the nature of the paramagnetic metal ion; those with longer electron relaxation times are the most effective on nuclear relaxation. The dependence on r^{-6} can be translated in distance restrains for structure calculations that help defining the relative position of the metal ion and protein nuclei. Measurement of intermolecular restraints exploiting paramagnetic probes attached to either of the two proteins in a complex is therefore an useful source of structural information. Of course the contact term, which is confined to the paramagnetic metal ion ligands, in general will not affect the intermolecular NMR parameters, while the dipolar terms will represent important sources of information. While for a given paramagnetic ion, the observed PRE only depends upon the distance (through a factor r^{-6}) between the measured nucleus on a protein and the paramagnetic center on the partner protein in the complex, PCS provides both distance (r^{-3}) and orientational information. Given the large magnetic moment of unpaired electron(s) (the magnetic moment of a single unpaired electron is about 600 times larger than the largest nuclear magnetic moment, i.e., ^{1}H), PRE effects are strong, may be operative at very large distances (up to 30 Å) (Fig. 4.5) and provide a very sensitive tool for detecting the presence of low population transient

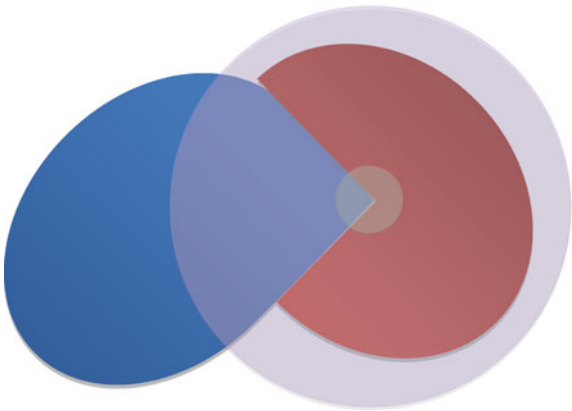


Fig. 4.5 Cartoon representation of a protein–protein complex. ^{1}H–^{1}H NOE extends up to 5 Å (*small transparent circle*), whereas PRE is larger range, extending up to 30 Å (*bigger circle*)

species in solution [31]. The observed intermolecular PRE relaxation rates in the fast exchange regime are population averages of all complexes present in solution and have been used to define the lowly populated states in protein complexes.

4.3.5 Weak and Ultra-Weak PPIs Online

NMR is an important structural method in biology. Since the first solution structure of a protein was solved by K. Wüthrich and coworkers in 1984 [32], about 9,000 NMR structures of proteins, nucleic acids and their adducts have been deposited in the Protein Data Bank (PDB), which amounts to about 12 % of the total number of structures in the PDB. About 9 % of the total NMR-based structures in the PDB are protein–protein and protein–nucleic acid complexes; under the category of protein–protein adducts, also protein-peptides complexes are counted. Presently, PDB only accepts structures, that is, depositions that are substantially determined by experimental data. This is the case of intermolecular complexes where structural restraints exist to define the relative orientation of the two interacting molecules. This class is clearly distinct from theoretical models, where the structural data are mainly determined by *in silico* approaches; they include, for example, homology modeling, but also data-driven docking procedures, such as those used in building structural models of complexes on the basis of chemical shift perturbation mapping. Theoretical models that have been previously released will continue to be publicly available via the existing models archive at ftp://wwpdb.org/pub/pdb/data/structures/models/current/. Another interesting website where one might recruit annotated structural information of proteins is that of Proteopedia (http://proteopedia.org/wiki/) and would be desirable that its use becomes a common practice by the authors of structural models of weak and ultra-weak PPIs.

4.3.6 Weak and Ultra-Weak Complexes by NMR: A Few Case Examples

A plethora of model structures based on chemical shift perturbation mapping data is available for protein–protein adducts of any strength [9, 33]. On the contrary, as discussed in the previous sections, the structural determination of weak/ultra-weak protein complexes is challenging and dominated by the NMR approach.

One example of NOE-based structures of ultra-weak PPIs is represented by the NMR structure of the complex between Nck-2 SH3-3 domain and PINCH-1 LIM4 domain, which has a K_d of $\sim 3 \times 10^{-3}$ M [34]. NOESY spectra acquired either on ^{13}C, ^{15}N-labeled LIM4 mixed with a slight excess of unlabeled SH3-3 or ^{13}C, ^{15}N-labeled SH3-3 mixed with a 4-fold excess of unlabeled LIM4 detected a total of 31 experimental intermolecular NOE constraints, whose assignment was fully

consistent with the chemical shift mapping data. The structure determination was performed in two steps: (a) calculations of the individual structures of SH3-3 and LIM4 in their bound forms; (b) calculations of the complex by using the bound SH3-3 and LIM4 structures as starting templates and by including the intermolecular NOEs. These two steps resulted in a well-defined structure of the complex with a polar and ultra-small interface (total buried surface area of 480 $Å^2$ almost equally distributed between the two domains) (PDB code: 1U5S).

When this structure was published, no other structures were available for complexes with $K_d > 10^{-3}$ M and only very few with $K_d > 10^{-4}$ M [35–37]. To extend applicability of NMR to this type of systems, the use of RDCs as additional orientational constraints in complex determination has been pursued by Blackledge and coworkers [38]. In 2009, the structure of the weakly interacting CD2AP (CD2-associated protein) SH3-C/ubiquitin complex using RDCs was published; it has a K_d of 132 μM, as determined from the simultaneous fit of the chemical shift perturbation of the most affected residues. While RDCs may become vital constraints when NOEs are hardly detected, as often happens in weak PPIs, their use is seriously compromised by the low population of the complex with respect to the isolated proteins. Indeed, under fast exchange conditions, measured RDCs report on both bound and free forms of the molecule. Nevertheless, their use was successfully applied in the structure determination of this complex thanks to a successful protocol including some key steps: (1) the RDCs of the free states are measured directly; (2) a titration is performed where RDCs are measured in different equilibrium mixtures of the free and bound form; (3) the RDCs originating from the bound state are obtained indirectly by extrapolation of the RDCs in the different equilibrium mixtures. Accurate measurements in the mixture were possible thanks to the combination of different isotopic labeling scheme of the partner proteins: mixtures of ^{15}N-labeled ubiquitin and ^{13}C, ^{15}N-labeled SH3-C allowed the measurement of N–H_N, Cα–Hα, Cα–CO, and CO–HN RDCs for SH3-C in HNCO and HN(CO)CA type experiments, while N–H_N RDCs were obtained for the single-labeled ubiquitin from a pair of spin-state-selected spectra where the signals of the double-labeled protein were suppressed. This type of RDC measurements with spectroscopic filtering permits to derive data for both partners in a given equilibrium mixture under identical alignment conditions, thus increasing precision in data comparison. The RDCs obtained by extrapolation to the fully bound form of ubiquitin and SH3-C were used to determine the relative orientation of the two proteins in the complex.

The homoeostasis of metal ions in cells occurs via a number of cellular pathways based on transient and often weak PPIs [39], characterized by small and polar interfaces. Metal transfer typically implies the formation of adducts where the metal itself acts as a bridge between proteins, by coordinating amino acids on both interacting partners. In principle, these metal mediated interactions would be particularly suitable for a characterization via paramagnetic restraints based on contact effects that would “enlighten” the resonances of the metal ligands. In practice, this approach has not yet been pursued, as the intracellular metal ion trafficking routes studied by NMR are essentially focused on diamagnetic cations,

with a very deep characterization of the systems involving copper(I) [40, 41]. Disappearance of paramagnetic effects on the iron axial ligands in the hemophore protein HasA upon formation of a stable complex with the receptor HasR has provided a strong spectroscopic evidence of the transfer of the heme cofactor from holoHasA to apoHasR upon establishment of the intermolecular adduct [24].

At variance with this, paramagnetic effects based on electron–nuclear interactions that are dipolar in origin have been largely pursued. Pioneering works focused on the characterization of electron transfer complexes. A nonfunctional complex between ferricytochrome b_5 and ferricytochrome c was studied using long-range inter-protein paramagnetic dipolar shifts [42]. Heteronuclear NMR spectra of samples containing ^{15}N-labeled cytochrome b_5 in complex with unlabeled cytochrome c allowed unambiguous assessment of pseudocontact shifts relative to diamagnetic reference states. As pseudocontact shifts in low spin iron(III) proteins, such as cytochromes, extend as much as 20 Å from the paramagnetic center, this approach is suitable for electron transfer proteins in fast exchange. Use of PCSs has been also applied to functional complexes like, for example, that of cytochrome f with plastocyanin ($K_d \sim 10^{-3}/10^{-2}$) [43, 44]. In electron transfer complexes, four different combinations for the oxidation states of the partner proteins are possible; the most relevant for the definition of the electron transfer mechanism is the one that corresponds to that of the reactants in the electron transfer reactions. However, given the high rate of this type of reactions, this state cannot be studied directly because it rapidly evolves toward the products. In practice, NMR has been applied to learn about the interaction of two electron transfer partners either using nonfunctional combination of oxidation states or using metal substitution for one of the partners. An example of the latter situation is the case of cytochrome f with plastocyanin from the Cyanobacterium Nostoc sp. PCC 7119 where, to characterize the oxidized and reduced complexes, two samples of ^{15}N-labeled plastocyanin and ^{15}N-labeled Cd-substituted plastocyanin were titrated with cytochrome f; the titration was followed with HSQC experiments [45]. From the comparison of complexes with oxidized or reduced cytochrome F was clear that many plastocyanin amide nuclei experience an additional shift caused by intermolecular PCSs from the ferric heme iron onto plastocyanin. The strongest PCSs were found in the hydrophobic patch region, suggesting that this area comes closest to the cytochrome F heme iron in the complex; charge complementarities were found between the rims of the respective recognition sites of cytochrome f and plastocyanin. At variance with this, in the complex of cytochrome f with plastocyanin from *P. laminosum* observed PCSs were small and did not result in a converged structure, strongly suggesting that this complex is more dynamic in nature. The data suggested an orientation of plastocyanin in which only the hydrophobic patch is in contact with cytochrome f, without a charge–charge interaction, in contrast to the other structures [46].

The idea of electron transfer complexes as extremely dynamic entities has been further pursued by Ubbink and coworkers using PRE and has lead to the description of the electron transfer complex formation as a multistep process, where the active complex is in equilibrium with the “encounter complex” [47, 48].

In the latter, the proteins do not assume a single orientation relative to each other, but rather they rapidly change orientation, thus sampling an ample part of the surface of the partner. The functional significance of the encounter complex shall be that of accelerating the formation of the specific active complex by reduction of the dimensionality of the conformational search. In the encounter complex, intermolecular interactions are dominated by long-range electrostatic forces that prolong the lifetime of the encounter and allow for pre-orientation of protein molecules, thus limiting the conformational search to a part of the binding surface. In the active complex, the proteins assume a single, well-defined orientation that is stabilized not only by electrostatic forces but also by short-range interactions like hydrogen bonds and hydrophobic interactions. Some ET complexes appear to be entirely or mostly nonspecific, with the encounter complex being the dominant form [49–56], whereas in others the specific complex dominates [57–59]. An interesting example of transient but specific structure determination by PRE is represented by the structure determination of pseudoazurin (Paz)/nitrite reductase (NiR) adduct solved in 2008 [57]. Multi-copper NiR catalyzes the reaction of nitrite reduction to nitric oxide after receiving electrons from a blue copper protein known as pseudoazurin. To determine the binding interface of Paz with NiR, a titration of $^{2}H–^{15}N$ Zn substituted Paz was performed with increasing amount of NiR and followed via TROSY spectra. The chemical shift mapping shows that the interaction area is situated in a very well-defined region, and the large chemical shift perturbations observed for some amino acids suggest that the complex should be in a single orientation. Fitting the chemical shift differences during the titration for a few residues of Paz a K_d of 64 μM was estimated. To determine the orientation of Paz relative to NiR, a rigid Gd probe was attached to two engineered Cys residues on NiR surface (^{Gd}NiR). The PRE measurement of Paz amide protons was performed on the $^{2}H–^{15}N$ labeled Paz–^{Gd}NiR paramagnetic complex and on the $^{2}H–^{15}N$ labeled Paz-^{Y}NiR diamagnetic one. The unpaired electron in the paramagnetic center of the Paz–^{Gd}NiR complex gives rise to relaxation effects in the range of 20–30 Å from the metal that are not present in the diamagnetic analogue. The chemical shift changes and the decreased intensities of the amide proton peaks of Paz in proximity of the gadolinium probe, converted into distances, were used as experimental information in the docking calculation. The docking of the two proteins resulted in a cluster of low energy structures with an average backbone rmsd of 1.5 Å. The binding region is confined in a small nonpolar areas centered on the exit/entry ports of the electron, surrounded by complementary charged residues. The interprotein Cu–Cu distance of about 15.5 Å allows for fast interprotein electron transfer. Conjugation with a paramagnetic label of CcP has allowed the characterization of the encounter complex between cytochrome c and CcP, where nearly equally populated encounter and specific complexes have been found and interpreted as a consequence of the biological function, striking the right balance between fast ET (that requires specificity) and a high turnover rate (that requires fast dissociation) [60]. PRE approaches have found applications also in unraveling the mechanistic details of association pathways between proteins involved in processes different from electron transfer [61, 62].

4.4 Protein–Protein Interactions Druggability

The ubiquitous nature and central role of PPIs make them very attractive targets for therapeutic intervention. However, they are characterized by large interaction areas: the size of the interface, defined as the variation in surface accessible area between the complex and the isolated proteins, is <1,500 $Å^2$ for transient complexes [63]. These areas tend to be quite flat, thus making more challenging the design of a potential small ligand molecule, because flatness limits the number of intermolecular noncovalent interactions that can be established. The above considerations have long suggested that a small molecule drug will be unable to substitute the partner biomolecule. Two factors contribute to lower this chemical risk [64]: (1) the existence of energetic hot spots at the interface and (2) the site adaptability of surface patches.

1. Hot spots are localized regions centered inside the broad contact area that account for most of the binding free energy in PPIs [65, 66]. Residues at hot spots are generally more conservative than the others and may be both hydrophobic and polar in nature. They are surrounded by energetically less important residues that also serve to exclude bulk solvent from the hot spots [67]. Small molecules targeting hot spots have the ability to inhibit PPIs and therefore to modulate biological activity.

The use of peptides as therapeutic agents is limited, given their low oral bioavailability, quick degradation and poor pharmacokinetics. Still, identification of hot spots may proceed using short peptides from contacted areas: if they show high affinity, they can represent valuable starting points for the development of peptidomimetics and small molecules. The contact surface of interacting proteins is often formed by different secondary structure elements, so that it may be difficult to design a homologous continuous bonding peptide. Nevertheless, peptides not sharing any sequence identity or homology with the natural partner protein have been sometimes found to posses very large affinities for protein surfaces. NMR has been very often used to structurally characterize the adducts between a target protein and any of these peptides [68–71] (Fig. 4.6).

2. Flexibility of the residues at the interfaces [72] makes them more druggable because the dynamics in the surface of the free protein will offer conformations that are more prone to binding. A structured free protein is adequately described as an ensemble of energetically similar conformers that interconvert. Induced fit, conformer selection and induction are different ways through which this protein flexibility translates into binding ability [73]. In the induced fit model, the ligand binds to the lowest energy conformation of the protein, which is then distorted to accommodate the ligand. Conformer selection is a process in which the ligand selectively binds to one of the preexisting conformers of the dynamic ensemble, increasing its population with respect to the others. Conformational induction describes a process through which ligand binding converts the protein into a conformation that it is not populated in its free state. The availability of

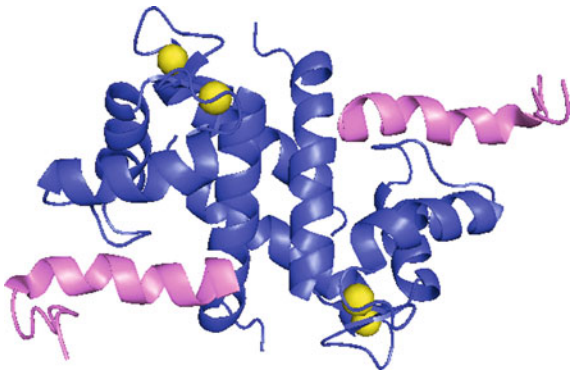


Fig. 4.6 NMR-determined structure of the calcium loaded S100B, from *Rattus norvegicus* (*blue ribbon*), bound to the negative regulatory domain of p53 (*magenta ribbon*). PDB ID: 1DT7

methods to describe proteins as dynamic ensembles rather than as rigid structures plays, therefore, a key role in drug discovery projects aimed at targeting PPIs. NMR is one of these methods. Structure determination of proteins by NMR results in a bundle of conformers equally consistent with the data; this bundle describes at least a subset of the conformational space that would be consistent with the experimental restraints, and therefore might represent a better starting point for the virtual screening methodologies that are widely used as a complementary approach to experimental screening in many drug discovery and development programs to enhance the probability of success at the stage of lead identification [74–76]. In principle, characterization of protein dynamics based on NMR relaxation measurements might further contribute to define the flexibility of the target protein [77, 78]. However, it should be noted that usually NMR dynamics studies are restricted to the protein backbone, while side chain conformational flexibility is probably the most relevant for the purposes of describing surface adaptability.

4.5 Identification of Inhibitors of Protein–Protein Interactions

The idea that small compounds preventing or modulating PPIs belonging to different classes of targets may have therapeutic potential has received great impulse during the last 10–15 years [64]. Several NMR techniques are available as powerful tools for the identification of active substances and for the understanding of the structural determinants of the binding process [12, 79, 80]. Here, the most common NMR experiments will be reviewed and critically analyzed with respect to their range of applicability and limitations. They span from medium-throughput approaches based on monodimensional ^{1}H NMR spectra, which allow for the efficient screening of the binding ability of libraries of hundreds of compounds, to two-dimensional ^{1}H–^{15}N HSQC spectra for the mapping of the protein binding

surface and eventually three-dimensional spectra for the direct observation of intermolecular constraints. Finally, novel tools to derive structural information from molecular docking will be briefly introduced.

4.5.1 Thermodynamics and Kinetics of the Binding of a Small Ligand to a Protein

As discussed in Sect. 4.2 for PPIs, binding affinity together with association and dissociation rate constants are relevant parameters for NMR spectroscopy binding experiments. The binding of a small ligand to a target protein can generally be described as a one-step reaction. In the diffusion limit approach, k_{on} values for the binding of small ligands are of the order of 10^7–10^8 s^{-1} M^{-1} [12].

Practically, when dealing with a small molecule ligand (MW < 1,000 Da) and a protein, we have fast exchange regime on the chemical shift time scale for $K_d > 10^{-6}$ M. For ligands with K_d in the 10^{-6}–10^{-7} M range, the timescales correspond to an intermediate exchange situation, usually resulting in line broadening during titration. Finally, for ligands with $K_d < 10^{-8}$ M, the slow exchange limit is commonly observed, in which two separate lines are present during the titration representing the free and bound forms of the molecule [81].

4.5.2 NMR for the Identification of Protein Ligands

One major step in drug discovery processes comprises the identification of substances with binding activity for the target protein and usually requires the screening of very extended libraries of compounds (from hundreds to thousands). NMR plays an important role for the identification of ligands at the screening level, using methods based on the observation of the signals of the small molecule that is present in solution in a large excess (80–100 fold) with respect to the target protein, used in its unlabeled form. In saturation transfer difference (STD) [82, 83] spectroscopy, the difference between a monodimensional ^{1}H NMR spectrum obtained by saturation of the protein signals and another spectrum identical to the first one, but without protein saturation is recorded. In the difference spectrum, only interacting molecules have nonzero intensity signals. The observed resonances are those of the free ligand: upon saturation of protein signals, the resonances of the bound ligand become saturated; by exchange between the free and bound forms, the saturation is carried out into the free form that is detected. The Water-Ligand Observed via Gradient SpectroscopY (W-LOGSY) [84] technique relies on an indirect excitation of the receptor-ligand complex through selective perturbation of the bulk water magnetization. Binders and nonbinders have differential cross-relaxation properties with water. Binders interact with water spins via dipolar interactions with sufficiently long rotational correlation times to yield negative cross-relaxation rates. By contrast, the dipolar interactions of nonbinders

with water have much shorter correlation time values, leading to positive cross-relaxation rates. The W-LOGSY, being based on the interaction between the protein and the solvent and on the different sign of the cross-correlation effects related to the rotational correlation time of the protein, is more suitable for polar and relatively large (>100 amino acids) systems with respect to STD. Both methods are based on the existence of a fast equilibrium between the bound and free ligand and therefore are not able to detect strongly binding ligands (i.e., those with $K_d < 10^{-5}$) with slow dissociation rates. However, this is not an issue at the initial stages of screening processes, when strong binders are not expected.

The above-described 1D NMR experiments are typically fast (5–10 min) and may use mixtures of reciprocally not-interacting and not-reacting small molecules. Ligand pooling allows for hundreds to thousands of compounds to be screened in a single day. If the presence of active molecules is observed, the experiments can repeated on the single ligands to identify the binder(s). Another advantage of ligand-based NMR methods is the minimal concentration of protein required ($\sim$20–30 μM) for each experiment. Additionally, isotopically labeled proteins are not needed for the NMR ligand affinity screening, and protein molecular weight is not a limiting factor: higher molecular weight proteins enhance the observation of a binding event in a ligand-based NMR screening. It should be kept in mind, however, that compounds with poor solubility are difficult to detect since the method requires the observation of the free ligand signals.

The diffusion coefficient of a small molecule ligand in solution is drastically altered upon binding to a large partner. Diffusion Ordered SpectroscopY (DOSY) [85] is suited to determine diffusion coefficients of molecules and is therefore useful for detecting association of ligands with proteins. The measured diffusion coefficient is the average of the diffusion coefficients of the free and bound ligand weighted by their molar fraction and an excess of ligand translates into little contribution of the bound form to the measured diffusion coefficient. Still, relatively high absolute concentrations of the ligand are needed to make the signals detectable in DOSY experiments. Therefore, at variance with STD and W-LOGSY, the best results are obtained at high protein concentrations (of the order of mM), and therefore, relatively large amounts of soluble protein are necessary for this technique.

Ligand-based NMR screening methods do not provide any structural information about the protein–ligand complex nor ensure the specificity of the binding to the targeted surface area on the partner protein. For these purposes, protein-based approaches are needed, which focus on changes in the protein NMR spectrum to identify a binding event. As for PPIs, when the three-dimensional structure of the protein is known and the sequence-specific NMR assignment of the backbone amide resonances has been obtained, ^{1}H–^{15}N HSQC-based chemical shift perturbation mapping [16] is used to study protein–small molecules interactions. It allows the identification of the residues at the binding area on the basis of the observed chemical shift perturbations as a function of the residue number (Garret plot) [15] (Fig. 4.7); in this case, the number of protein residues involved is lower than in the case of PPIs because the surface interaction area is smaller with respect

to the protein–protein adducts. $^{1}H–^{15}N$ TROSY, CRINEPT_TROSY or CRIPT–TROSY experiments are used for the same purposes for large protein targets (see Sect. 4.3.3).

4.5.3 Determination of Ligand Affinity by NMR

After having identified ligands that specifically interact with the targeted protein surface area, the estimate of the binding affinity constitutes another important piece of information. A number of biophysical methods are able to provide *in vitro*

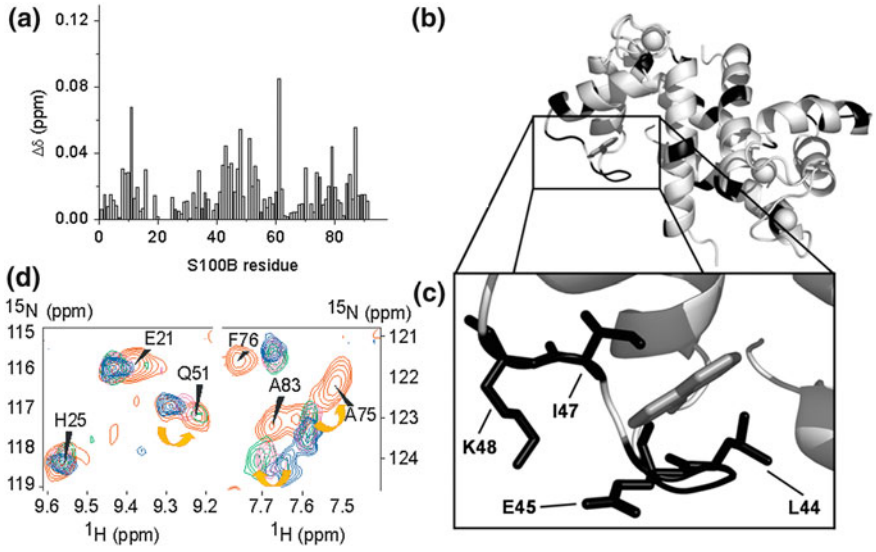


Fig. 4.7 Addressing hot spots for the interaction between S100B and p53: binding site characterization by chemical shift mapping using as case example the S100B/1-naphthol adduct. **a** The Garrett plot reports the weighted average chemical shift change for ^{1}H and ^{15}N between the initial and the final stage of a titration as a function of the residue number. **b** Chemical shift mapping of residues (highlighted in *black*) experiencing significant chemical shift variations ($\Delta\delta > 0.03$ ppm) on the structure of the S100B dimer; **c** Close-up of the S100B/1-naphthol interaction: in *black* sticks are represented the residues with non-negligible chemical shift perturbation closest to the ligand. **d** Competition binding experiments: superimposition of two areas of the $^{1}H–^{15}N$–HSQC spectra of i) S100B/1-naphthol adduct at 1:2 ratio (*red*); ii) S100B/p53 at 1:2 ratio (*blue*); iii) S100B/p53/1-naphthol at 1:2:2 ratio (*magenta*); iv) S100B/p53/1-naphthol at 1:2:4 (*green*). From the comparison of the four spectra, it results that the addition of the 1-naphthol fragment to the S100B-p53 peptide complex has different effects on the various peaks. The peaks that in the ternary complex return to the same shift they had in the S100B-1-naphthol adduct binary complex identify relevant hot spots on the protein, where the 1-naphthol binds and displaces the p53 peptide. From the analysis of the chemical shift mapping is deduced that 1-naphthol replaces a portion of the bound peptide p53 from S100B binding. The other portion of the p53 peptide was proposed to rearrange in a new conformation. Adapted from Ref. [98]

K_d values (namely surface plasmon resonance, fluorescence, calorimetry). As far as NMR is concerned, K_ds can be derived from both ligand-based experiments and target-based experiments. The latter set is the same as that described for PPIs in Sect. 4.3.1: to determine the K_d for complex formation, NMR spectra are recorded at each point in a titration of one ligand into a solution of the protein.

On the other hand, ligand-based competition binding experiments have been used to test specificity of an identified ligand and to extract, with titration experiments, the dissociation binding constant [86]. The initial step of the approach requires the identification of a medium- to low-affinity ligand and the determination of its binding constant (for example by ITC measurements). The attenuation of the W-LOGSY signal of the reference compound in the presence of a competitor depends on the relative value of the dissociation constants for the reference compound and the competitor.

Competition binding experiments extend the ability of NMR methods to detect high affinity ligands, using a known low-affinity ligand as an indicator. A significant reduction or disappearance of the signals of the indicator in W-LOGSY or STD experiments proves the presence of a high affinity ligand, if it competes with the indicator for the same binding site.

4.5.4 Definition of the Ligand Binding Mode

In fragment-based projects, when a weak binder (or preferred scaffold) is found to bind to a target protein, follow-up hit optimization strategies can be devised to iteratively increase the affinity of the compound. Based on its ability to provide structural information, NMR has been extensively validated not only as an efficient technique for the initial screening and identification of weakly interacting fragments but also for subsequently guiding their optimization into molecules with higher affinity and more favorable drug-like properties. Knowledge of the structural determinants of the affinity of the small molecule for the protein surface of interest constitutes essential information for any optimization protocol. In general, docking programs are able to identify the correct binding site from the chemical shift perturbation mapping on the protein, but leave open the problem of the ligand orientation, unless intermolecular NOEs can be detected [87], eventually with heteronuclear filtered approaches [88, 89]. Indeed, methods have been established to identify intermolecular NOEs between labeled proteins and unlabeled peptides and/or small ligands, while omitting signals from intramolecular NOEs within both labeled and unlabeled constituents. Combined with the structure of the target (if available), NOEs data could result in the structure of the complex between the fragment ligand and the target. With this information in hand, it is possible to apply the principles of structure-based drug design to mature the initial fragment into a higher affinity ligand. In the absence of intermolecular NOEs, data-driven docking approaches can be used. There are several programs [80] where NMR chemical shift perturbations from a protein–ligand binding interaction are first

qualitatively used to guide molecular docking by defining a search grid on the protein surface to generate a corresponding structure and then to filter the docking quantitatively: the docked model is validated by an agreement with the experimental chemical shift perturbations observed for the protein. Energy-based molecular docking programs generate different poses that can be scored to define which best represents the experimental conformation using, at least in principle, a filter based on data different from protein chemical shift perturbations. As mentioned above, given the three-dimensional structure of a protein, the binding pose of a ligand can be determined using distance restraints derived from assigned protein–ligand NOEs. The main limitation of this approach is the need for resonance assignments of the ligand-bound protein that usually do not simply involve the backbone amides, but extend to the side chains of the residues at the interface. An approach has been proposed that utilizes NOE data without requiring protein NMR resonance assignments; only the ^{1}H NMR assignments of the bound ligand are essential [90]. For each trial binding pose, the 3D NOESY spectrum is predicted, and the predicted and observed patterns of protein–ligand NOEs are matched and scored, thus identifying the best scoring poses.

4.5.5 A Case History

This case history has been selected to supply the reader with an example of a successful application of NMR in drug discovery: here small organic molecules that bind to proximal subsites of a protein are identified, optimized and linked together to produce high affinity ligands. With this technique, compounds with nanomolar affinities for the binding protein were discovered by tethering two ligands with micromolar affinities through appropriate linkers.

Proteins of the B cell lymphoma 2 (Bcl-2) family represent important regulators of apoptotic cell death [91, 92]. Proapoptotic members such as Bax, BaK, Bid and Bad can interact with antiapoptotic members such as Bcl-2 and Bcl-Xl, giving rise to heterodimers responsible for the inhibition of apoptosis [92, 93]. For example, Bcl-2 and Bcl-Xl proteins inhibit apoptosis by binding pro-apoptotic BAK and BAD molecules; the inhibition of the interaction of Bcl-2/Bcl-Xl proteins with Bak/Bad molecules represents a target in the treatment of different cancers. Three regions (BH1, BH2 and BH3) of the antiapoptotic members of family participate in the binding of proapoptotic members, whereas only the BH3 binding site of the proapoptotic members has been found to play a critical role. Small truncated peptides from the BH3 region of Bak (16 mer) and Bad (25 mer) have been found necessary and sufficient for promoting the biological activity and binding to Bcl-Xl. The NMR solution structures of Bcl-Xl bound to Bak and Bad BH3 peptides have shown that the three BH domains of the antiapoptotic protein form an hydrophobic cleft in which the peptides bind [92, 93]. Small molecules able to bind at this BH3 binding site should be able to inhibit this interaction. An initial set of inhibitors targeting Bcl-Xl was identified using a fragment-based methods and

parallel synthesis to obtain molecules that bind in the hydrophobic groove of Bcl-Xl [94]. A first-site ligand, a fluorobiarylacid, with K_d of about 300 μM was identified through an NMR-based screen of a 10,000-compound fragment library, with an average MW of 210 Da. Chemical shift perturbation mapping based on $^{1}H–^{15}N$ HSQC spectra allowed the identification of the area of interaction at the center of the hydrophobic groove of Bcl-Xl, approximately at the same position where a key Leu of Bak peptide binds; the comparison of the structure–activity relationships of several biaryl analogues permits to determine that the carboxylate group present on this binder is critical for the interaction. Another key interaction between Bak peptide and Bcl-Xl involves Ile85 on the former, which identifies a second hot spot on the protein. To identify molecules interacting in this second site, another screening was carried out using a library of 3,500 compounds, with an average MW of 125 Da, in the presence of an excess of the fluorobiarylacid molecule. The identified interacting molecules were several naphthol analogues presenting K_d values in the low mM range. The NMR solution structure of the ternary complex of Bcl-Xl with these two binders was solved and used to develop an appropriate linking strategy. Various linkers were used, and one molecule with K_d of the order of 1.4×10^{-6} M was identified. Nine intermolecular NOEs allowed the determination of a model structure from which a better linker was built, and the solution structure of the adduct of Bcl-Xl with this new binder was determined and used for further optimization of the second binding site ligand. The final molecule had a K_d of 3.6×10^{-8} M. The molecule needed further optimization to increase its solubility and reduce its affinity for albumin. To reduce binding to human serum albumin (HSA), a structure-based approach was used [95]. The NMR structure of an analogue of this ligand complexed with domain III of HSA was compared to the structure of the ligand bound to Bcl-Xl. The structures revealed portions of the molecule that were solvent exposed in the Bcl-Xl complex, but surrounded by lipophilic residues in the complex with albumin. These positions were modified with polar substituents to reduce binding to albumin without affecting affinity for Bcl-Xl. The resulting compound, called ABT-737, binds with high affinity (K_d 0.6×10^{-9} M) to Bcl-Xl, and its nanomolar activity is retained in the presence of 10 % HSA. Further pursuing this approach, Abbott Laboratories have developed a potent, orally bioavailable inhibitor (ABT-263) that shows robust antitumor activity *in vivo*, but has severe side effects like thrombocytopenia, which has been shown to be mediated by inhibition of Bcl-Xl, but not by that of Bcl-2. Therefore, in 2010, the use of NMR and structure-based drug design in the discovery of selected inhibitors of Bcl-2 was published [96]. A library of about 17,000 compounds with an average MW of 225 Da was screened through the NMR. At variance with the common chemical shift mapping of backbone amide resonances in $^{1}H–^{15}N$ HSQC spectra, compound binding was monitored by following chemical shift perturbations of Leu, Val and Ile methyl groups in a $^{1}H–^{13}C$–HSQC spectrum. A diphenylmethane has been discovered that binds to Bcl-2 with a K_d of 20×10^{-6} M and is 20-fold more selective for Bcl-2 than for Bcl-Xl. Fifteen intermolecular NOEs allowed to obtain a structural model of the adduct. In analogy with the strategy developed for Bcl-Xl, ternary

complexes have been characterized and NMR structural data used to design an appropriate link. The theoretical gain in affinity for the linking of two compounds is the product of their dissociation constants plus a contribution from the linker. At the moment, linking of a compound with K_d of 20×10^{-6} M to another with K_d of 400×10^{-6} M has produced a derivative with a K_d of 220×10^{-9} M, that is, almost 30 times higher than expected. This is most probably attributable to a distorted binding of the original compound due to some strain induced by the linker. A final selective inhibitor is still under study.

References

1. Rowe AJ (2011) Ultra-weak reversible protein–protein interactions. Methods 54:157–166
2. Schreiber G, Haran G (2009) Zhou HX (2009) Fundamental aspects of protein–protein association kinetics. Chem Rev 109:839–860
3. Alsallaq R, Zhou HX (2008) Electrostatic rate enhancement and transient complex of protein–protein association. Proteins 71:320–335
4. Archakov AI, Govorun VM, Dubanov AV, Ivanov YD, Veselovsky AV, Lewi P, Janssen P (2003) Protein–protein interactions as a target for drugs in proteomics. Proteomics 3:380–391
5. Prudencio M, Ubbink M (2004) Transient complexes of redox proteins: structural and dynamic details from NMR studies. J Mol Recognit 17:524–539
6. Corzo J (2006) Time, the forgotten dimension of ligand binding teaching. Biochem Mol Biol Educ 34:413–416
7. Jensen MR, Ortega-Roldan JL, Salmon L, van Nuland N, Blackledge M (2011) Characterizing weak protein–protein complexes by NMR residual dipolar couplings. Eur Biophys J 40:1371–1381
8. Takeuchi K, Wagner G (2006) NMR studies of protein interactions. Curr Opin Struct Biol 16:109–117
9. Bonvin AM, Boelens R, Kaptein R (2005) NMR analysis of protein interactions. Curr Opin Chem Biol 9:501–508
10. Zuiderweg ER (2002) Mapping protein–protein interactions in solution by NMR spectroscopy. Biochemistry 41:1–7
11. Crowley PB, Ubbink M (2003) Close encounters of the transient kind: protein interactions in the photosynthetic redox chain investigated by NMR spectroscopy. Acc Chem Res 36:723–730
12. Meyer B, Peters T (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew Chem Int Ed 42:864–890
13. Zerbe O, Mannhold R, Kubinyi H, Folkers G (2003) BioNMR in drug research. Wiley-VCH, Zurich
14. Vaynberg J, Qin J (2006) Weak protein–protein interactions as probed by NMR spectroscopy. Trends Biotechnol 24:22–27
15. Garrett DS, Seok YJ, Peterkofsky A, Clore GM, Gronenborn AM (1997) Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the escherichia coli phosphotransferase system. Biochemistry 36:4393–4398
16. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW (1996) Discovering high-affinity ligands for proteins: SAR by NMR. Science 274:1531–1534
17. de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 5:883–897
18. Dominguez C, Boelens R, Bonvin AM (2003) HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125:1731–1737

19. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T_2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94:12366–12371
20. Riek R, Wider G, Pervushin K, Wüthrich K (1999) Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. Proc Natl Acad Sci USA 96:4918–4923
21. Fiaux J, Bertelsen EB, Horwich AL, Wüthrich K (2002) NMR analysis of a 900 K GroEL GroES complex. Nature 418:207–211
22. Griswold IJ, Dahlquist FW (2002) Bigger is better: megadalton protein NMR in solution. Nat Struct Biol 9:567–568
23. Horst R, Bertelsen EB, Fiaux J, Wider G, Horwich AL, Wütrich K (2005) Direct NMR observation of a substrate protein bound to the chaperonin. GroEL. Proc Natl Acad Sci USA 102: 12748-12753
24. Caillet-Saguy C, Piccioli M, Turano P, Izadi-Pruneyre N, Delepierre M, Bertini I, Lecroisey A (2009) Mapping the Interaction between the Hemophore HasA and its outer membrane receptor has R using CRINEPT-TROSY NMR spectroscopy. J Am Chem Soc 131: 1736–1744
25. Debye PJW (1929) Polar molecules. Dover Publications Inc
26. Ikura M, Bax A (1992) Isotope-filtered 2D NMR of protein-peptide complex: study of Skeletal muscle myosin light chain kinase. J Am Chem Soc 114:2433–2440
27. Banci L, Bertini I, Cefaro C, Cenacchi L, Ciofi-Baffoni S, Felli IC, Gallo A, Gonnelli L, Luchinat E, Sideris DP, Tokatlidis K (2010) Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import. Proc Natl Acad Sci USA 107:20190–20195
28. Prestegard JH, Bougault CM, Kishore AI (2004) Residual dipolar couplings in structure determination of biomolecules. Chem Rev 104:3519–3540
29. Capozzi F, Casadei F, Luchinat C (2006) EF-hand protein dynamics and evolution of calcium signal transduction: an NMR view. J Biol Inorg Chem 11:949–962
30. Bertini I, Calderone V, Cerofolini L, Fragai M, Geraldes CFGC, Hermann P, Luchinat C, Parigi G, Teixeira JMC (2012) The catalytic domain of MMP-1 studied through tagged lanthanides. FEBS Lett 586:557–567
31. Volkov AN, Ubbink M, Van Nuland NAJ (2010) Mapping the encounter state of a transient protein complex by PRE NMR spectroscopy. J Biomol NMR 48:225–236
32. Williamson MP, Marion D, Wuthrich K (1984) Secondary structure in the solution conformation of the proteinase inhibitor IIA from bull seminal plasma by nuclear magnetic resonance. J Mol Biol 173:341–359
33. O'Connell MR, Gamsjaeger R, Mackay JP (2009) The structural analysis of protein–protein interactions by NMR spectroscopy. Proteomics 9:5224–5232
34. Vaynberg J, Fukuda T, Chen K, Vinogradova O, Velyvis A, Tu Y, Ng L, Wu C, Qin J (2005) Structure of an ultraweak protein–protein complex and its crucial role in regulation of cell morphology and motility. Mol Cell 17:513–523
35. Wang JH, Meijers R, Xiong Y, Liu JH, Sakihama T, Zhang R, Joachimiak A, Reinherz EL (2001) Crystal structure of the human CD4N-terminal two-domain fragment complexed to a class II MHC molecule. Proc Natl Acad Sci USA 98:10799–10804
36. Kang RS, Daniels CM, Francis SA, Shih SC, Salerno WJ, Hicke L, Radhakrishnan I (2003) Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding. Cell 113:621–630
37. Sundquist WI, Schubert HL, Kelly BN, Hill GC, Holton JM, Hill CP (2004) Ubiquitin recognition by the human TSG101 protein. Mol Cell 13:783–789
38. Ortega-Roldan JL, Jensen MR, Brutscher B, Azuaga AI, Blackledge M, van Nuland NA (2009) Accurate characterization of weak macromolecular interactions by titration of NMR residual dipolar couplings: application to the CD2AP SH3-C: ubiquitin complex. Nucleic Acids Res 37:e70

39. Banci L, Bertini I, Calderone V, Della Malva N, Felli IC, Neri S, Pavelkova A, Rosato A (2009) Copper(I)-mediated protein–protein interactions result from suboptimal interaction surfaces. Biochem J 422:37–42
40. Banci L, Bertini I, McGreevy KS, Rosato A (2010) Molecular recognition in copper trafficking. Nat Prod Rep 27:695–710
41. Banci L, Bertini I, Cantini F, Ciofi-Baffoni S (2010) Cellular copper distribution: a mechanistic systems biology approach. Cell Mol Life Sci 67:2563–2589
42. Guiles RD, Sarma S, DiGate RJ, Banville D, Basus VJ, Kuntz ID, Waskell L (1996) Pseudocontact shifts used in the restraint of the solution structures of electron transfer complexes. Nat Struct Biol 3:333–339
43. Ubbink M, Lian LY, Modi S, Evans PA, Bendall DS (1996) Analysis of the ^{1}H-NMR chemical shifts of Cu(I)−, Cu(II)- and Cd-substituted pea plastocyanin. Metal-dependent differences in the hydrogen-bond network around the copper site. Eur J Biochem 242:132–147
44. Hulsker R, Baranova MV, Bullerjahn GS, Ubbink M (2008) Dynamics in the transient complex of plastocyanin-cytochrome f from Prochlorothrix hollandica. J Am Chem Soc 130:1985–1991
45. Diaz-Moreno I, Diaz-Quintana A, De la Rosa MA, Ubbink M (2005) Structure of the complex between plastocyanin and cytochrome f from the cyanobacterium nostoc Sp. PCC 7119 as determined by paramagnetic NMR. J Biol Chem 280:18908–18915
46. Bashir Q, Scanu S, Ubbink M (2011) Dynamics in electron transfer protein complexes. FEBS J 278:1391–1400
47. Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein–protein association. Nature 444:383–386
48. Volkov AN, Worrall JAR, Holtzmann E, Ubbink M (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. Proc Natl Acad Sci USA 103:18945–18950
49. Xu X, Reinle W, Hannemann F, Konarev PV, Svergun DI, Bernhardt R, Ubbink M (2008) Dynamics in a pure encounter complex of two proteins studied by solution scattering and paramagnetic NMR spectroscopy. J Am Chem Soc 130:6395–6403
50. Liang ZX, Nocek JM, Huang K, Hayes RT, Kurnikov IV, Beratan DN, Hoffman BM (2002) Dynamic docking and electron transfer between Zn-myoglobin and cytochrome b(5). J Am Chem Soc 124:6849–6859
51. Volkov AN, Ferrari D, Worrall JA, Bonvin AM, Ubbink M (2005) The orientations of cytochrome c in the highly dynamic complex with cytochrome b5 visualized by NMR and docking using HADDOCK. Protein Sci 14:799–811
52. Liang ZX, Kurnikov IV, Nocek JM, Mauk AG, Beratan DN, Hoffman BM (2004) Dynamic docking and electron-transfer between cytochrome b5 and a suite of myoglobin surface-charge mutants. Introduction of a functional-docking algorithm for protein–protein complexes. J Am Chem Soc 126:2785–2798
53. Worrall JA, Liu A, Crowley PB, Nocek JM, Hoffman BM, Ubbink M (2002) Myoglobin and cytochrome b5: a nuclear magnetic resonance study of a highly dynamic protein complex. Biochemistry 41:11721–11730
54. Worrall JA, Reinle W, Bernhardt R, Ubbink M (2003) Transient protein interactions studied by NMR spectroscopy: the case of cytochrome C and adrenodoxin. Biochemistry 42:7068–7076
55. Hoffman BM, Celis LM, Cull DA, Patel AD, Seifert JL, Wheeler KE, Wang J, Yao J, Kurnikov IV, Nocek JM (2005) Differential influence of dynamic processes on forward and reverse electron transfer across a protein–protein interface. Proc Natl Acad Sci USA 102:3564–3569
56. Ubbink M, Bendall DS (1997) Complex of plastocyanin and cytochrome c characterized by NMR chemical shift analysis. Biochemistry 36:6326–6335

57. Vlasie MD, Fernández-Busnadiego R, Prudêncio M, Ubbink M (2008) Conformation of pseudoazurin in the 152 kDa electron transfer complex with nitrite reductase determined by paramagnetic NMR. J Mol Biol 375:1405–1415
58. Ubbink M, Ejdebaeck M, Karlsson BG, Bendall DS (1998) The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic NMR and restrained rigid-body molecular dynamics. Structure 6:323–335
59. Volkov AN, Worrall JAR, Holtzmann E, Ubbink M (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. Proc Natl Acad Sci USA 103:18945–18950
60. Bashir Q, Volkov AN, Ullmann GM, Ubbink M (2010) Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome *c* and cytochrome *c* peroxidase. J Am Chem Soc 132:241–247
61. Fawzi NL, Doucleff M, Suh JY, Clore GM (2010) Mechanistic details of a protein–protein association pathway revealed by paramagnetic relaxation enhancement titration measurements. Proc Natl Acad Sci USA 107:1379–1384
62. Villareal VA, Spirig T, Robson SA, Liu M, Lei B, Clubb RT (2011) Transient weak protein–protein complexes transfer heme across the cell wall of Staphylococcus aureus. J Am Chem Soc 133:14176–14179
63. Nooren IMA, Thornton JM (2003) Structural characterisation and functional significance of transient protein–protein interactions. J Mol Biol 325:991–1018
64. Veselovsky AV, Archakov AI (2007) Inhibitors of protein–protein interactions as potential drugs. Curr Comput: Aided Drug Des 3:51–58
65. Arkin MR, Wells JA (2004) Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. Nat Rev Drug Discov 3:301–317
66. Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. Science 267:383–386
67. Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280:1–9
68. Agamennone M, Cesari L, Lalli D, Turlizzi E, Del Conte R, Turano P, Mangani S, Padova A (2010) Fragmenting the S100B–p53 interaction—Combined virtual/biophysical screening approaches to identify ligands. Chem Med Chem 5:428–435
69. Rustandi RR, Baldisseri DM, Weber DJ (2000) Structure of the negative regulatory domain of p53 bound to S100B (betabeta). Nat Struct Biol 7:570–574
70. Inman KG, Yang R, Rustandi RR, Miller KE, Baldisseri DM, Weber DJ (2002) Solution NMR structure of S100B bound to the high-affinity target peptide TRTK-12. J Mol Biol 324:1003–1014
71. Ivanenkov VV, Jamieson GA Jr, Gruenstein E, Dimlich RV (1995) Characterization of S-100b binding epitopes. Identification of a novel target, the actin capping protein, CapZ. J Biol Chem 270:14651–14658
72. Gaffen SL (2001) Signaling domains of the interleukin 2 receptor. Cytokine 14:63–77
73. Teague SJ (2003) Implications of protein flexibility for drug discovery. Nat Rev Drug Discov 2:527–541
74. Zhong S, Macias AT, Mackerell AD Jr (2007) Computational identification of inhibitors of protein–protein interactions. Curr Top Med Chem 7:63–82
75. Gonzalez-Ruiz D, Gohlke H (2006) Targeting protein–protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. Curr Med Chem 13:2607–2625
76. DeLano WL (2002) Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 12:14–20
77. Kay LE (1998) Protein dynamics from NMR. Nat Struct Biol 5:513–517
78. Ishima R, Torchia DA (2000) Protein dynamics from NMR. Nat Struct Biol 7:740–743
79. Pellecchia M, Bertini I, Cowburn D, Dalvit C, Giralt E, Jahnke W, James TL, Homans SW, Kessler H, Luchinat C, Meyer B, Oschkinat H, Peng J, Schwalbe H, Siegal G (2008) Perspectives on NMR in drug discovery: a technique comes of age. Nat Rev Drug Discov 7:738–745

80. Stark JL, Powers R (2011) Application of NMR and molecular docking in structure-based drug discovery. Top Curr Chem. doi:10.1007/128_2011_213
81. Fesik SW, Zuiderweg ER, Olejniczak ET, Gampe RT Jr (1990) NMR methods for determining the structures of enzyme/inhibitor complexes as an aid in drug design. Biochem Pharmacol 40:161–167
82. Meinecke R, Meyer B (2001) Determination of the binding specificity of an integral membrane protein by saturation transfer difference NMR: RGD peptide ligands binding to integrin alpha$_{IIb}$beta$_3$. J Med Chem 44:3059-3065
83. Wang YS, Liu D, Wyss DF (2004) Competition STD NMR for the detection of high-affinity ligands and NMR-based screening. Magn Reson Chem 42:485–489
84. Dalvit C, Fogliatto G, Stewart A, Veronesi M, Stockman BJ (2001) Water LOGSY as a method for primary NMR screening: practical aspects and range of applicability. J Biomol NMR 21:349–359
85. Wu D, Chen A, Johnson CS (1995) An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. J Magn Reson A 115:260–264
86. Dalvit C, Fasolini M, Flocco M, Knapp S, Pevarello P, Veronesi M (2002) NMR-based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands. J Med Chem 45:2610–2614
87. Assfalg M, Bertini I, Del Conte R, Giachetti A, Turano P (2007) Cytochrome c and organic molecules: the solution structure of the para-aminophenol adduct. Biochemistry 46:6232–6238
88. Bertini I, Calderone V, Cosenza M, Fragai M, Lee Y-M, Luchinat C, Mangani S, Terni B, Turano P (2005) Conformational variability of MMPs: beyond a single 3D structure. Proc Natl Acad Sci USA 102:5334–5339
89. Isaksson J, Nystroem S, Derbishire W, Wallberg H, Agback T, Kovacs H, Bertini I, Giachetti A, Luchinat C (2009) Does a fast nuclear magnetic resonance spectroscopy- and X-ray crystallography hybrid approach provide reliable structural information of ligand-protein complexes? A case study of metalloproteinases. J Med Chem 52:1712–1722
90. Constantine KL, Davis ME, Metzler WJ, Mueller L, Claus BL (2006) Protein-ligand NOE matching: a high-throughput method for binding pose evaluation that does not require protein NMR resonance assignments. J Am Chem Soc 128:7252–7263
91. Adams JM, Cory S (1998) The Bcl-2 protein family: arbiters of cell survival. Science 281:1322–1326
92. Sattler M, Liang H, Nettesheim D, Meadows RP, Harlan JE, Eberstadt M, Yoon HS, Shuker SB, Chang BS, Minn AJ, Thompson CB, Fesik SW (1997) Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. Science 275:983–986
93. Petros AM, Nettesheim DG, Wang Y, Olejniczak ET, Meadows RP, Mack J, Swift K, Matayoshi ED, Zhang H, Thompson CB, Fesik SW (2000) Rationale for Bcl-xL/Bad peptide complex formation from structure, mutagenesis, and biophysical studies. Protein Sci 9:2528–2534
94. Petros AM, Dinges J, Augeri DJ, Baumeister SA, Betebenner DA, Bures MG, Elmore SW, Hajduk PJ, Joseph MK, Landis SK, Nettesheim DG, Rosenberg SH, Shen W, Thomas S, Wang X, Zanze I, Zhang H, Fesik SW (2006) Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. J Med Chem 49:656–663
95. Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, Rosenberg SH (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature 435:677–681
96. Petros AM, Huth JR, Oost T, Park CM, Ding H, Wang X, Zhang H, Nimmer P, Mendoza R, Sun C, Mack J, Walter K, Dorwin S, Gramling E, Ladror U, Rosenberg SH, Elmore SW, Fesik SW, Hajduk PJ (2010) Discovery of a potent and selective Bcl-2 inhibitor using SAR by NMR. Bioorg Med Chem Lett 20:6587–6591

97. Bertini I, Chevance S, Del Conte R, Lalli D, Turano P (2011) The anti-apoptotic Bcl-xL protein, a new piece in the puzzle of cytochrome c interactome. PLoS ONE 6:e18329
98. Arendt Y, Bhaumik A, Del Conte R, Luchinat C, Mori M, Porcu M (2007) Fragment docking to S100 proteins reveals a wide diversity of weak interaction sites. Chem Med Chem 2:1648–1654

Chapter 5
Protein–Protein Interactions in the Solid State: The Troubles of Crystallizing Protein–Protein Complexes

Stefano Mangani

5.1 Introduction

The term "protein–protein interface" (PPI) encompasses very different macromolecular assemblies:

- Oligomeric/multimeric proteins made of identical (homologous) subunits.
- Oligomeric/multimeric proteins made of different (heterologous) subunits.

All the above components may be competently assembled to perform specific functions (e.g., oligomeric enzymes) or may lead to pathogenic species by either gain of function (GOF), loss of function (LOF), or both as exemplified by proteins involved in neurological disorders like α-synuclein and the prion protein PrP^C [1] or by the classical sickle cell hemoglobin mutant [2].

The characterization of the type of intermolecular interactions that lead to the formation of a functionally competent PPI is probably the most relevant step to understand why it is formed, how, and whether it might be disrupted and possibly regulated. Many experimental tools are available for performing the task of revealing the structure and the nature of protein–protein interactions and their power and number are rapidly increasing as many frontier techniques are becoming powerful and widely available (see a list in Ref. [3]).

The determination of the tri-dimensional structure of PPIs is the main step for every research program aimed to the discovery of molecules able to interfere with protein–protein interactions. The main issue regarding the structure determination by X-ray crystallography of PPIs and of the interactions with molecules able to interfere or regulate such interactions deals with the ability to obtain such complexes in the crystalline state and not on the crystal structure determination on itself. This chapter will mainly treat the experimental problems related to

S. Mangani (✉)
Department of Biotechnology, Chemistry and Pharmacy, University of Siena,
Via Aldo Moro 2, 53100 Siena, Italy
e-mail: stefano.mangani@unisi.it

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4_5,

obtaining pure proteins and protein complexes and their crystallization strategies. Selected examples from the literature are provided.

5.2 Oligomeric Homologous Proteins

Oligomeric proteins or enzymes made of identical subunits (homologous) probably represent the best candidates for successful structural studies both in terms of easiness of cloning, expression, purification, and in terms of crystallization [4]. Heterologous oligomers are less common and may present more difficult cases (see Sect. 5.3). Overall, the Protein Data bank contains (at the time of writing) about 13,400 non-redundant oligomeric homologous protein structures and 2,500 oligomeric heterologous protein structures determined by single-crystal X-ray diffraction.

However, even for those cases that appear the most favorable, laboratory practice tells that the road to obtain pure, controlled, samples for structural studies is paved with difficulties and failures. Only a fraction of the successfully expressed and purified protein reaches the state of complete structure determination, as indicated by the overall success ratio (NMR or X-ray data) of structural genomics consortia of 6.9 %, with respect to the cloned proteins [5].

The quaternary structure of oligomeric proteins is very often characterized by symmetry. Indeed, the majority of soluble and membrane-bound proteins are symmetrical oligomeric complexes made of two or more subunits. Hypotheses have been formulated about the evolutionary conservation of symmetrical multimers [6, 7]. The evolutionary selection of symmetrical oligomeric complexes is driven by functional, genetic, and physicochemical needs. Large proteins are selected for specific morphological functions, such as formation of rings, containers, and filaments, and for cooperative functions, such as allosteric regulation and multivalent binding. Large proteins are also more stable against denaturation and have a reduced surface area exposed to solvent when compared with many individual, smaller proteins. Large proteins are constructed from oligomeric building blocks for reasons of error control in synthesis, coding efficiency, and regulation of assembly. Finally, symmetrical oligomers assembled in closed forms may be favored because of stability and the lower probability to undergo uncontrolled polymerization, misfolding, and formation of fibrils [8].

The quaternary structure of a protein is not easily established from the crystallographic structure. Several approaches have been used to rationalize protein assemblies and to predict their oligomeric state from structural data, resulting in extremely useful research tools like the PISA Web server [7, 9–13].

It is even more difficult to establish the correct quaternary structure of an oligomeric protein in vivo, as the dissociation equilibria between monomers and the different possible oligomerization states depend on the chemical environment that, in turn, influences the protein function. Consequently, depending on local pH, monomer concentration, ionic force, concentration of ligands, redox potential, etc.,

the same protein might assemble in different quaternary structures that might differ from cell compartment to cell compartment.

The different cell parts (cytoplasm, nucleus, Golgi, mitochondria, etc.) are quite crowded environments as the total macromolecule concentration in a cell is of the order of 200–300 gL^{-1} [14–16] and a variety of molecules, beside proteins, are present and might promote or disfavor protein–protein interactions. The total particle concentration in a cell is of the order 10^{-1} M (100–200 mM), and many species, like metal ions and/or anions such as chloride, bicarbonate, might influence the aggregation state of proteins. Furthermore, the availability of chemical species is strictly regulated by control systems present in the cell that avoid the release of potentially dangerous chemicals. This is the case of metal ions that are essential for life as structural components of proteins, as catalytic cofactors of hydrolytic and redox enzymes, as part of electron transfer processes in the cell [17], but are potentially toxic if present in an improper form and quantity. In order to provide the cell with the right amount of the appropriate metal ion, an extensive regulatory and protein-coding machinery has evolved to maintain the homeostasis of the required metal ions and to regulate the metal uptake, efflux, and intracellular trafficking. Such machinery involves several, species-specific, protein systems that have also the role sequester the metal ion and to protect it from undesired reactions [18]. For example, the global cell concentration of a metal ion like Zn(II) is of the order of 0.1 mM, although the free form (aquoion) is in the picomolar range [19]. It has been shown that the formation of protein–protein complexes might be facilitated by this metal ion, as exemplified by the cases of tonin (PDB: 1TON) [20] and of a bacterial superoxide dismutase (PDB: 1S4I) [21], where the zinc ions present in the crystallization medium promote the formation of intermolecular complexes. In the latter case, the binding of the zinc ion also causes the transition from the unstructured form of superoxide dismutase to the properly folded protein as shown in Fig. 5.1.

Observations like the ones reported above might be considered artifacts of the crystallization procedure and the observed complexes considered as non-physiological aggregates. However, these results provide experimental evidence that such complexes might form in environments where the appropriate local concentration of molecules or ions, zinc ions in this case, is reached and that chemical equilibria involving water species (pH), small molecules or metal ions, and the protein govern the aggregation state of the macromolecule.

A recently discovered class of proteins, called morpheeins, is a paramount example that proteins may assemble in different quaternary structures and have different properties depending on the local environment. The prototype of this class is the enzyme porphobilinogen synthase that constitutes a new paradigm for allosteric control of enzyme function [22–24]. Morpheeins are homo-oligomeric proteins whose function is controlled through a reversible transition between different quaternary assemblies that depend on conformational variation of a limited portion of the protein, induced by changes in pH, protein concentration, or substrate concentration. Different morpheein oligomers present functional differences that may constitute a valuable starting point for drug discovery since the

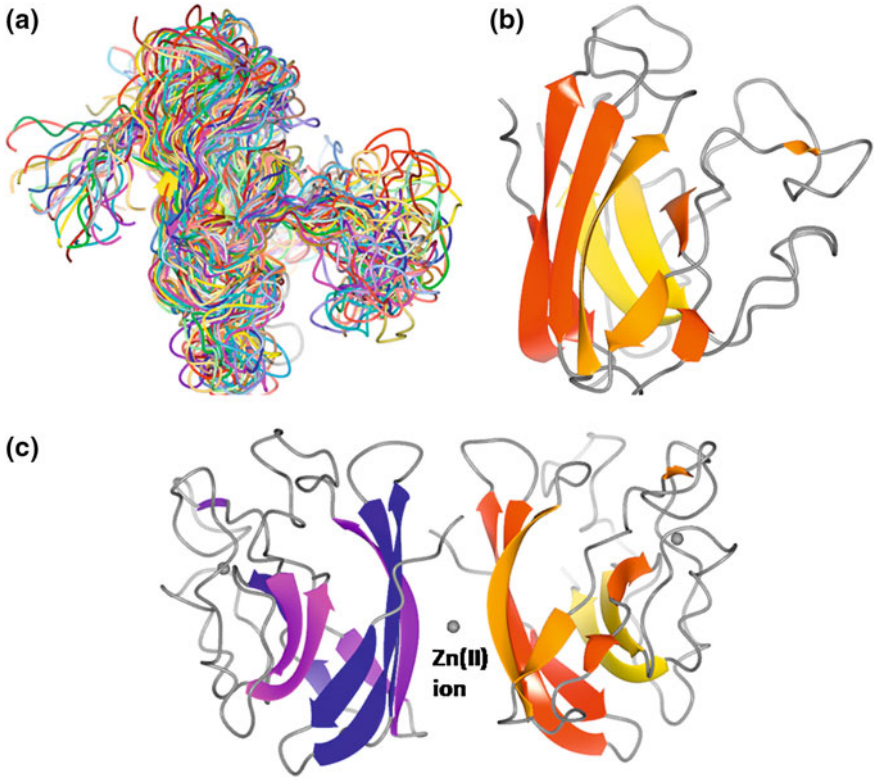


Fig. 5.1 **a** Bundle of NMR structures of BsSOD determined in solution (PDB code 1U3N). The protein appears in a molten globule state where the β-barrel supersecondary structure characteristic of superoxide dismutase is barely recognizable. **b** View of a subunit of BsSOD as determined by X-ray diffraction. The protein is completely structured and displays the usual superoxide dismutase fold. **c** The dimeric structure of BsSOD found in the crystal (PDB code 1S4I). The two BsSOD subunits are held together by intersubunit coordination bonds formed by a bridging zinc(II) ion and Asp and His residues occurring at the interface of each subunit [21]

morpheein function can be regulated by small molecules able to shift the equilibrium between the different forms, toward the inactive morpheein form, by preferentially binding one of them [24–26]. The occurrence of such mechanism of action has been documented for an inhibitor of porphobilinogen synthase [27]. Probably, morpheeins constitute a larger group of proteins than is supposed today. A recent list of proteins having the morpheein characteristics of allosteric regulation is presented in Ref. [18].

Another enlightening example of the influence of the chemical environment on the protein aggregation state is provided by the protein tubulin. Tubulin has a structural role in the cell and builds complex structures via extensive association of smaller building blocks. Tubulin builds up microtubules by assembling dimers of alpha- and beta- tubulin when GTP is bound to the molecule [28, 29]. Once GTP is

hydrolyzed to GDP, the microtubule disassembles [30, 31]. The structure of tubulin dimers has been determined by electron crystallography, and the structure has been used to fit the high-resolution electron microscopy image to reconstruct the whole polymer [32]. Later, the structure has also been determined by X-ray crystallography in complex with colchicine and with the stathmin-like domain (SLD) of RB3 showing the formation of a tubulin curved complex, which prevents its incorporation into microtubules halting the microtubule formation [33].

Further difficulties in establishing the functionally relevant protein interface arise when the protein complex formation is transient and involves weak free energy of binding. The work of Krissinel [13] has shown that small free energy of dissociation $\left(\Delta G^0_{diss} \leq 3 - 5 \text{ kcal mol}^{-1}\right)$ corresponds to a very high probability of such interactions being misrepresented in the crystallographic experiment. In other words, it is difficult to obtain information and observe the correct protein–protein interaction for such weak protein complexes in the crystal state. The PISA server appears to perform much better than crystallographic structure determination in identifying these weak interactions [13].

A search of the protein interaction thermodynamic (PINT) database[1] (http://www.bioinfodatabase.com/pint/index.html) [34] shows only four entries referring to crystal structures of protein complexes having a ΔG^0_{diss} in the range 3–5 kcal mol^{-1} (PDB codes: 2PCC, 2PCB, 1QG1, 1SPS). It should be noted, however, that the free energy of binding of the partner proteins strongly depends from the ionic strength of the medium with an inverse proportionality law. As an experimental confirmation of this, the crystallization conditions of all above entries report very low ionic strengths and PEGs or MPD as precipitating agents (BMCD http://xpdb.nist.gov:8060/BMCD4/index.faces; [35]). It is then possible that crystals obtained in low ionic strength will show the correct interactions. The above observations suggest that the crystallization screening of choice for protein–protein complexes with low free energy of dissociation should be PEGs/MPD screens at low ionic strength.

In any case, weak complexes constitute a relevant target for drug design, as they are involved in important pathways such as the intracellular signaling that, by its nature, has to be transient. Because all the above-mentioned difficulties, it is not surprising that only few cases of structural determination of weak complexes are available in the scientific literature and that most of the structural information

[1] The attempt to satisfy the evident need to couple structural data and thermodynamic parameters of protein–protein complexes has been undertaken by compiling the protein–protein interactions thermodynamic (PINT) database (http://www.bioinfodatabase.com/pint/index.html) [34]. PINT reports experimental data of protein–protein interactions such as dissociation constant (Kd), association constant (Ka), free energy change (ΔG), enthalpy change (ΔH), and heat capacity change (ΔCp) associated with the protein sequence and structure. Also, the experimental methods and the related literature are available. The PINT database is a unique tool as it provides the possibility to understand the factors that determine protein–protein interactions and their specificity.

about the binding of small molecule inhibitors of such weak interactions is obtained from binary protein-small molecule inhibitors.

For all the reasons mentioned above, crystallographic techniques alone cannot always provide the correct answer to the question of determining the functional quaternary structure of a protein. The experimentalist should use a variety of structural techniques operating at different resolution and in different environments. Ideally, high-resolution structural determinations (X-ray crystallography, NMR) should be coupled to techniques operating at the micro/nano-meter scale, like high-resolution cryo-electron microscopy [36–42], fluorescence microscopy [43], and/or to techniques operating in different conditions like X-ray small-angle scattering [44]. In special cases, when metal ion mediates protein–protein interactions, X-ray absorption spectroscopy (XAS), coupled with NMR or X-ray diffraction [45, 46], may be used to explore the conditions favoring the complex formation. An example is provided by the X-ray absorption study on the bacterial copper chaperone protein CopZ, where ionic strength and small molecule ligands like ascorbate or dithiothreitol were found to affect the formation of protein dimers [47]. The possibility of small molecules to favor and affect protein association has been recognized since a long time [48]. Nowadays, it is exploited to direct protein assembly in order to understand biochemical processes and to build "synthetic" biological systems [49].

5.2.1 Obligate Homo-Oligomeric Proteins

The obligate homo-oligomeric proteins constitute a further interesting class of proteins [50]. These are proteins complexes where the protomers are not found isolate in vivo and usually perform their physiological function only as multimers. This is the case of enzymes made of identical subunits, where the active site occurs at the subunit interface or where the activity strongly depends on the oligomeric state, and represent very interesting targets for designing a new kind of inhibitor molecules able to bind and/or disrupt the interface. Homodimeric repressors like the Met repressor homodimer provide examples of functionally obligate complexes. A list of some obligate multimeric enzymes is reported in a recent review by Cardinale et al. [51] where the subject of targeting this class of enzymes for inhibitors directed at the interface is deeply analyzed.

The cases where molecules directed toward the interfaces of homo-oligomeric enzymes have advantages over the active-site-directed inhibitors are exemplified by human thymidylate synthase (hTS). hTS is a homodimeric enzyme which is discussed in Chap. 2.

5.3 Oligomeric Heterologous Protein–Protein Complexes

The crystal studies performed on heterologous protein–protein complexes probably constitute the major achievements of X-ray crystallography and surely are milestones of our comprehension of the function at molecular level of the most relevant biological processes. To support this statement is sufficient to look at the few examples reported in Table 5.1. All the macromolecular complexes reported in the Table have been obtained as intact complexes (or, in some cases as a fraction of the complex) from the originating organism. This is due to the sufficiently high stability of the complex in the native state that allows purification procedures and crystallization trials. The idea of obtaining the macromolecular particles from extremophile organisms like *Haloarcula marismortui* and, to some extent, also *Tetrahymena thermophila* in the case of ribosomes, provides a clear example of the concept that macromolecular complexes from extremophiles must be characterized by stronger intermolecular interactions and hence higher stability and easiness of manipulation. However, in every case, all the steps of the crystal structure determination from purification to crystallization, phasing, and refinement have required years of efforts and great skills and intuitions to reach success.

The main obstacle in using the native organism for obtaining heterologous complexes arise from the fact that, in the vast majority of cases, the protein complexes exist in very low quantities in the natural host. Although highly efficient purification techniques such as tandem affinity purification of tagged open reading frames [52–54] are now at hand and widely used, they still provide very little quantities with respect to the amount required for crystallization experiments. Furthermore, as it is always the case when dealing with proteins purified from the natural host, the complexes may suffer from heterogeneity. In the case of high eukaryotes is quite common the occurrence of protein complexes that exist as mixtures of isoforms differing in subunit composition and/or containing post-translational modifications, resulting in a variety of species that prevent crystallization.

The alternative ways of obtaining the protein assemblies are as follows:

1. The heterologous expression of the whole complex (co-expression) in the same host cell.
2. The in vitro reconstitution of the complex starting from the single protein components expressed separately.

Important technological advances have been made allowing co-expression of recombinant proteins in both prokaryote and eukaryote host cells. Expression in eukaryotic cells, such as yeast, baculovirus expression in insect cells, or mammalian cell lines, has the advantage to preserve posttranslational modifications that might be essential for protein function and stability and also because these systems are equipped with chaperone proteins that may assure the proper protein folding. A list of successful production of co-expressed protein complexes is reported in a recent review [55]. The European SPINE2 initiative (http://www.spine2.eu/

Table 5.1 Some multiprotein complexes expressed and purified in the native state

Complex	Resolution	MW/Subunits	Preparation	Organism
Respiratory complex I (NADH dehydrogenase) [108]	3.0–6.3	1–3 MDa/15–40	Intact complex (partial or whole)	*T. thermophilus; E. coli; Y. lipolytica*
Respiratory complex II (succinate:ubiquinone oxidoreductase) [109]	2.4	124 kDa/4	Intact complex	Sus scrofa
Respiratory complex II (fumarate reductase) [110, 111]	2.2–3.0	120–131 kDa/4–3	Intact complex	*E.coli; W. succinogenes*
Respiratory complex III (cytochrome bc1 complex) [112]	3.0	267 kDa/11 + 1Cytc bound (1 half of the functional molecule)	Intact complex	Yeast
Cytochrome C oxidase [68]	1.8	205 kDa/13 (1 half of the functional molecule)	Intact complex	Bos taurus
Atp synthase F1 [113]	2.0	351 kDa/7	Intact complex	Bos taurus
Ribosome bacterial 30S [114]	3.0	783 kDa/20 + 1 RNA	Intact complex	*T. thermophilus*
Ribosome eukaryotic 40S [115]	3.9	1.21 MDa/34 + 1 RNA	Intact complex	*T. thermophila*
Ribosome bacterial 50S [116]	2.4	1.40 MDa/27 + 2 RNA	Intact complex	*H. marismortui*
Ribosome eukaryotic 60S [117]	3.5	2 MDa/44 + 2 RNA	Intact complex	*T. thermophila*
20S proteasome [118]	1.9	1.4 MDa/28	Intact complex	*S. cerevisiae*

SPINE2/) dedicated to the structural determination of protein complexes has provided new automated platforms for the production of complexes in mammalian and eukaryotic cell lines [56, 57].

On the other hand, eukaryotic expression systems present the disadvantage of being expensive, of requiring special sterile environments (insect and mammalian cells), slow growth and, sometimes, low yield. Furthermore, the produced proteins might be heavily glycosylated making difficult (or impossible) the crystallization or, in alternative, requiring deglycosylation procedures that introduce further problems, reduce the overall yield and may cause heterogeneity at the molecular level.

For all the above reasons, bacterial hosts (mainly *Escherichia coli*) remain the first choice for expressing protein complexes [58–60]. The advantages of expressing proteins or protein complexes in *E. coli* consist on having (in most cases, but not always) high yield at low cost and in short time. Both cloning and expression might be integrated and automatized as described for the high-throughput platforms realized in SPINE2 [57, 59]. Furthermore, *E. coli* expression ensures the absence of glycosylation and of other posttranslational modifications that represent an advantage for X-ray crystallography studies, where non-homogenous protein preparations prevent crystallization. A further advantage of co-expression is that the assembled complex is characterized by a higher solubility than the individual components alone [61].

A recent successful example of the co-expression of protein complexes strategy is provided by the crystal structure determination of the paraspeckle protein PSPC1-NONO heterodimer [62, 63]. However, crystals were obtained only after optimization of the lengths and of the termini of the heterodimer [62]. This is a quite common requirement in the design of a protein expression experiment as both N- and C-termini heavily influence the ability of the protein to pack properly in the crystal. Both bioinformatic resources and experimental techniques should be used to optimize the construct and to check the conformation of the proteins in solution before the crystallization experiments as described in a recent protocol [4].

Different considerations should be made for heterologous complexes that either are integral membrane proteins or have a membrane anchoring subunit. The traditional approach is to obtain the complexes from the native organism through purification of massive amount of the original membrane contents as exemplified by the classical work on ATP synthase [64–67], on cytochrome C oxidase [68], on bovine cytochrome bc1 complex [69] crystal structure determinations. This approach is still successfully pursued as indicated by the recent structure of the Na^+K^+ pump obtained from Sus scrofa [70]. However, huge efforts have been recently made to develop more efficient methods to produce crystallization grade membrane proteins by introducing engineered bacterial expression systems. Following the observation that although many membrane proteins can be overexpressed and obtained as inclusion bodies, they do not properly refold giving the functional protein [71], alternative methods to obtain membrane proteins were envisaged. Years ago, *E. coli* mutant strains able to overexpress membrane proteins have been isolated and successfully used (the “Walker strains”) [72]. Overexpression is paralleled by the proliferation of intracellular membranes in the

host, but such abundant products are often toxic for the cell leading to cell death [71].

Based on these premises, Wagner et al. [73] have studied membrane protein overexpression in the Walker strains and have been able to engineer the new *E. coli* Lemo21(DE3) strain, a BL21(DE3)-derived strain, that is tunable for overexpression by controlling the activity of the T7 RNA polymerase using its natural inhibitor T7 lysozyme (T7Lys). This strain allows optimization of the overexpression of both membrane and soluble proteins.

Besides *E. coli*, successful expression of membrane proteins has been obtained in yeast and high-throughput structural genomics platforms to maximize the success rate of producing crystallization quality have been designed [74].

A wide range of possible difficulties in obtaining a heterocomplex might be encountered depending on the nature of the target of study. Often, the complex is functionally transient and is not possible to isolate it in the host expression system. In other instances, the complex formation constant might be too low to allow co-purification. These problems might be overcome by using the strategy of expressing the partner proteins separately and then obtain the complex by mixing the purified proteins in the required stoichiometric ratio. Examples of this procedure are provided by the successful crystallization and crystal structure determination of protein–peptide complexes like pilotin-secretin [75] or the yeast Dyn2-Nup159 [76] where the peptides, constituting the binding fragment of the partner protein, have been synthesized and used to prepare the complex directly in the crystallization solution. A second successful example of reconstituting a complex by mixing the partner proteins in the proper amount is provided by the case of CuSBA complex [77]. Cus proteins are present in *E. coli* where they assemble an efflux system of the resistance–nodulation–division (RND) type that has the function to expel from the cell antibiotics and toxic inorganic cations like copper(I) and silver(I). Three proteins form a CusCBA system with formula $CusC_3CusB_6CusA_3$ where CusC forms an outer membrane channel, CusB is a membrane fusion protein, and CusA is the proton motive force-dependent inner membrane efflux pump. The co-crystals of the CusBA part of the system have been obtained by mixing the separately expressed and purified CusB and CusA proteins [78, 79]. Interestingly, the crystal structure of the CusBA complex shows the correct assembly with stoichiometry $CusB_6CusA_3$ (see Fig. 5.2), despite that the crystals were grown from a solution containing equimolar (0.1 mM CusA and 0.1 mM CusB in a buffer solution [77]) amounts of the two proteins. This fact indicates that despite the affinity of the two proteins is in the μM range [77], the interactions between the CusB and CusA partners are enough to drive the formation of the heterocomplex.

Other examples of functionally transient heterocomplexes are provided by the metal trafficking proteins that are responsible for shuttling metal ions within different compartments of the cell to finally assemble the functional enzyme that uses the metal cofactor for catalysis [80, 81]. These proteins form transient complexes by complementary shape and electrostatic recognition of the partner and by sharing coordination bonds to the transported metal ion with solvent-exposed

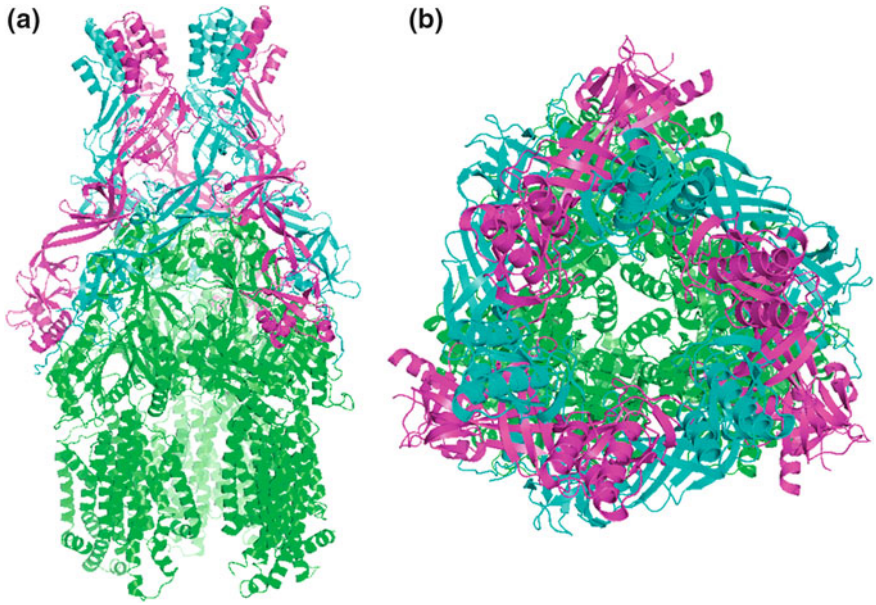


Fig. 5.2 Crystal structure of the $CusB_6CusA_3$ complex (PDB code:3NE5) [77]. CusB molecules are represented as *cyan and magenta ribbons* and CusA molecules are represented as *green ribbons*. **a** Side view. **b** Top view

ligands as in the case of the Atx1–Ccc2 complex [82]. Atx1 is a copper metallochaperone that supply Cu(I) ions to the Ccc2 copper ATPase that is able to export copper across a membrane [83].

To analyze the Atx1–Ccc2 interaction, the complex was formed in solution by mixing the separately expressed proteins and studied by NMR spectroscopy. The same approach has also been used to obtain crystals of the complexes between yeast superoxide dismutase and its copper chaperone CCS [84], the human copper(I)-chaperone HAH1 (human ATX1 homologue), and a metal-binding domain of the partner P-type copper-transporting ATPase [85].

A third type of preparing heterologous complexes for crystallization is a modification of the ones considered above and consists in cloning and purifying the proteins separately, then preparing the complex by mixing the partners in proper ratios, and finally purifying the required complex from the solution by chromatographic techniques. This approach has been successfully used in the crystallization of the complex between human transferrin and the *Neisseria meningitidis* TbpA-transferrin receptor used by *Neisseria* to extract iron directly from the human host [86].

The analysis of the literature, as well as the personal experience, tells that neither magic bullets nor safe recipes do exist for crystallizing difficult proteins or protein–protein complexes. All the examples reported in Table 5.1 are indicative of how the process of crystallization constitutes a paradigm of scientific discovery

on itself. In all instances, the crystallogenesis has posed great challenges that could be overcome only by careful observation of the protein behavior and inspired thoughts driven by scientific expertise. Here, I want to mention one of the most recent examples, which has contributed to awarding the 2012 Nobel Prize for Chemistry to the studies on G-coupled protein receptor (GCPRs). This refers to the crystallization of GPCRs in general [87–105] and in particular to the complex between the β2 adrenergic receptor (β2AR) and the stimulatory G protein that activates adenylyl cyclase (Gs) [106]. The crystallization story of the β2AR–Gs complex, described in Ref. [89], represents an enlightening example of the combined effect of chemical knowledge about the system under study, intuition, and enormous amount of experimental work needed to accomplish such achievement. Different types of experimental progresses had to be attained to reach the goal of crystallizing the complex:

(a) to find the right detergents;
(b) to use a high-affinity agonist to stabilize β2AR, and to realize that guanosine di/monophosphate (GDP/GMP) must be removed from Gs by using a pyrophosphatase, since both GDP and GMP prevent the high-affinity interaction between β2AR and Gs. Furthermore, the pyrophosphate analogue phosphonoformate (foscarnet) had to be used to stabilize the empty nucleotide binding pocket of Gs;
(c) the replacement of the unstructured amino terminus of the β2AR with T4 lysozyme [107] used to stabilize the receptor;
(d) finally, the extensive use of single-particle electron microscopy to check the conformational status of the complex at all stages.

5.4 Crystallization of Protein Complexes

Protein crystallization is a science on its own and crystallizing either proteins, protein–protein homo- or heterocomplexes, protein–nucleic acid, or protein–small molecule complexes requires careful planning, chemical and physical knowledge of the system to be crystallized, and patient recording of experimental observations in order to achieve success [4]. Anyone who is involved in protein crystallization knows that obtaining crystals of sufficient diffraction quality is always an unpredictable event, despite the progresses made in the field both on the theoretical understanding of the phenomenon and in terms of the large toolbox available [119–121]. In light of the above issues, only some general considerations can be made with respect to this difficult and essential step in the process of crystal structure determination. Usually, crystallization of heterologous protein–protein complexes presents higher difficulties with respect to the case of homologous ones because weaker intermolecular interactions exist, on average, in the former, as well as a tendency toward a limited solubility [50, 122–124]. As a rule of thumb,

the chances of crystallization are inversely proportional to the size and directly proportional to the formation constant of the complex.

The analysis of the crystallization conditions of the protein–protein complexes present in the PDB allowed Radaev et al. [124], Radaev and Sun [125] to establish that the distributions of the crystallization parameters for such complexes are more restricted than those of uncomplexed proteins allowing the proposal of dedicated crystallization screening kits specific for protein–protein complexes.

The crystal structure determination of protein-small molecule inhibitors of the protein–protein interactions appears to be even more difficult. Analysis of the available data shows that very few examples are available, despite the extensive research activity dedicated to discovery of PPI inhibitors over the last 10 years [123, 126, 127]. A structural database of protein–protein interaction inhibitors (2P2I database) has been recently developed (http://2p2idb.cnrs-mrs.fr) [128, 129], which contains (at the time of writing) 44 entries involving 42 different ligands distributed within 11 families of PPI (for example, 11 2P2I entries are related to inhibitors of Bcl-xL/Bak [130] interaction and 10 to inhibitors of the XIAP_BIR3/Smac [131] interaction). The quite small size of the 2P2I database, compared to the number of complexes present in the Protein Data Bank and the 33 entries of the 2P2I database determined by crystallographic methods, clearly indicate, in the first instance, the difficulty of finding inhibitors of PPIs and, secondly, the difficulty of obtaining crystals suitable for diffraction measurements.

One example of the difficulties encountered in the crystallization of protein complexes with inhibitors directed at disrupting PPIs can be taken from our laboratory and refers to the crystallization of the complex between S100B and one inhibitor of the S100B–p53 interaction [132]. The inhibitor, identified through bioinformatic tools and NMR screening of a library of compounds [132], is a hydrophobic molecule soluble in organic solvent and available as a racemic mixture. The hydrophobicity of the molecule imposes to use the co-crystallization approach in a water/dimethylsulfoxide (DMSO) mixture and to screen a large variety of conditions where, besides the usual parameters (protein concentration, pH, precipitant type, and concentration), also the inhibitor concentration and the water/DMSO ratio must be varied. In addition, the use of the organic solvent has the negative effect to decrease the crystal ordering and hence the quality of the diffraction data. The crystals of S100B in complex with the interface (surface) ligand SEN205A were eventually obtained after a lengthy optimization of the crystallization conditions and the structure solved [132].

Further examples of small molecule binary complexes with partners of protein–protein interactions are provided by a SMAC-mimetic compound bound to the XIAP regulator protein BIR3 domain [133], of a pro-death protein BH3 mimic that binds the Bcl-xL pro-survival protein [134], just to mention a few.

As indicated by the examples above, the job of obtaining a crystal of a complex with a small molecule inhibitor at the interface presents a variety of difficulties when the crystallographic study has the goal to find the binding mode of lead molecules obtained from a library preliminary screen. In this case, the Kd of the complex is usually quite high (in the micromolar range) and the interaction at the

interface (surface) is non-optimal. In these conditions, poor binding and/or multiple conformations of the ligand are likely to occur that result in electron density maps of difficult interpretation or, in the worst case, in complete absence of interpretable electron density.

Nevertheless, several crystal structure determinations of small molecules bound at interfaces where contacts with protein partners exist have been published. The 2P2I database holds a list of crystal structures with small molecule interface inhibitors bound to validated targets [129].

Even more challenging is the task to obtain inhibitors that bind at the interfaces of constitutive oligomeric proteins. Among these, the crystal structure of the binary complex between tumor necrosis factor (TNF-α) and the inhibitor obtained by linking a dimethylamine spacer a trifluoromethylphenyl indole and dimethyl chromone moieties is of particular interest [135]. The binding of the inhibitor prevents the binding of TNF-α to its receptor tumor necrosis factor receptor 1 (TNFR1) and the crystal structure of the complex readily explains this behavior. The compound had displaced one of the subunits from the TNF-α trimer resulting in a TNF-α dimer separated by the third subunit, all of them retaining the same basic structural subunit fold. This subunit arrangement of TNF-α prevents its binding to the TNFR1 receptor. The crystal structure, together with the kinetics of subunit dissociation obtained from a fluorescence homoquenching-based assay, indicates that the inhibitor binds to the intact trimeric TNF-α at the interface between two subunits by exploiting breathing movements of the molecule and causing subunit dissociation [135].

A confirmation of the above mechanism has been very recently reported [136]. The work describes the crystal structure of the binary complex between the cytokine CD40 ligand of the TNF family (CD40L) and the synthetic compound BIO8898 that inhibits CD40L binding to the CD40 membrane-bound receptor. The crystal structure CD40L revealed that one BIO8898 inhibitor molecule binds per CD40L trimer. Similarly to the previous case, the compound binds not at the surface of the protein, but intercalates between two subunits of the homotrimeric cytokine, disrupting a constitutive PPI and breaking the protein's 3-fold symmetry. However, BIO8898 does not displace one of the three cytokine subunits, but only distorts the CD40L trimer in such a way as to avoid the receptor binding. The authors [136] point out that these results show that it is possible to disrupt constitutive PPI characterized by very high stability and suggest that this possibility relates to the hydrophobic character of such constitutive interfaces [122] that allows hydrophobic molecules to intercalate. The challenge remains to be able to design small molecule inhibitors of constitutive PPIs that have the potency, selectivity, and pharmacological properties required by drug candidates.

Besides the above examples, other cases of inhibitors binding at the interface of constitutive oligomers are quite rare in the literature. One of them is the X-ray structure of an inhibitor peptide bound to the interface of the hTS dimer [137] which is described in Chap. 2.

The fact that PDB holds a wealth of information about not only protein–protein interactions, but also in terms of small molecules/ions binding sites possibly

located at the protein interface has been recently recognized and exploited in the compilation of the IBIS database.

The IBIS database (http://www.ncbi.nlm.nih.gov/Structure/ibis/ibis.html) [138, 139] enables researchers to conveniently study biomolecular interactions that have been observed in protein structures and, through inference by homology, to formulate predictions/hypotheses for biomolecular interactions, even if the data for specific biomolecules are not available. Therefore, IBIS can be considered a resource for functional annotation of proteins that have relevant homologues in the PDB.

The tools reviewed in this chapter are extremely useful to proceed further in the difficult task of obtaining structural information on PPI to be applied to the discovery of new molecules able to interfere with them. The few examples reported in the literature of PPI inhibitors either natural or synthetic and the even more rare cases of molecules that have progressed to clinical trials [140–147] should stimulate more efforts in this field.

References

1. Winklhofer KF, Tatzelt J, Haass C (2008) The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. EMBO J 27:336–349
2. Allison AC (2009) Genetic control of resistance to human malaria. Curr Opin Immunol 21:499–505
3. Perrakis A, Musacchio A, Cusack S, Petosa C (2011) Investigating a macromolecular complex: the toolkit of methods. J Struct Biol 175:106–112
4. Benvenuti M, Mangani S (2007) Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. Nat Protoc 2:1633–1651
5. Terwilliger TC, Stuart D, Yokoyama S (2009) Lessons from structural genomics. Ann Rev Biophys 38:371–383
6. Goodsell DS, Olson AJ (2000) Structural symmetry and protein function. Annu Rev Biophys Biomol Struct 29:105–153
7. Henrick K, Thornton JM (1998) PQS: a protein quaternary structure file server. Trends Biochem Sci 23:358–361
8. Taylor JP, Hardy J, Fischbeck KH (2002) Toxic proteins in neurodegenerative disease. Science 296:1991–1995
9. Ponstingl H, Henrick K, Thornton JM (2000) Discriminating between homodimeric and monomeric proteins in the crystalline state. Proteins 41:47–57
10. Ponstingl H, Kabir T, Thornton JM (2003) Automatic inference of protein quaternary structure from crystals. J Appl Crystallogr 36:1116–1122
11. Krissinel E, Henrick K (2007) Protein interfaces, surfaces and assemblies service PISA at European Bioinformatics Institute, http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html
12. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797
13. Krissinel E (2011) Macromolecular complexes in crystals and solutions. Acta Crystallogr D Biol Crystallogr 67:376–385
14. Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu Rev Biophys Biomol Struct 22:27–65
15. Minton AP (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J Biol Chem 276:10577–10580

16. Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597–604
17. Bertini I, Gray HB, Stiefel E, Valentine JS (2007) Biological inorganic chemistry structure and reactivity University Science Books. Sausalito, CA
18. Tottey S, Harvie DR, Robinson NJ (2005) Understanding how cells allocate metals using metal sensors and metallochaperones. Acc Chem Res 38:775–783
19. Maret W (2009) Molecular aspects of human cellular zinc homeostasis: redox control of zinc potentials and zinc signals. Biometals 22:149–157
20. Fujinaga M, James MN (1987) Rat submaxillary gland serine protease, tonin: structure solution and refinement at 1.8 Å resolution. J Mol Biol 195(2):373–396
21. Banci L, Bertini I, Calderone V, Cramaro F, Del Conte R, Fantoni A, Mangani S, Quattrone A, Viezzoli MS (2005) A prokaryotic superoxide dismutase paralog lacking two Cu ligands: from largely unstructured in solution to ordered in the crystal. Proc Natl Acad Sci USA 102:7541–7546
22. Jaffe EK (2005) Morpheeins: a new structural paradigm for allosteric regulation. Trends Biochem Sci 30:490–497
23. Jaffe EK (2010) Morpheeins: a new pathway for allosteric drug discovery. Open Conf Proc J 1:1–6
24. Jaffe EK, Lawrence SH (2012) The morpheein model of allostery: evaluating proteins as potential morpheeins. Methods Mol Biol 796:217–231
25. Lawrence SH, Jaffe EK (2008) Expanding the concepts in protein structure-function relationships and enzyme kinetics: teaching using morpheeins. Biochem Mol Biol Educ 36:274–283
26. Selwood T, Jaffe EK (2012) Dynamic dissociating homo-oligomers and the control of protein function. Arch Biochem Biophys 519:131–143
27. Lawrence SH, Ramirez UD, Tang L, Fazliyez F, Kundrat L, Markham GD, Jaffe EK (2008) Shape shifting leads to small-molecule allosteric drug discovery. Chem Biol 15:586–596
28. Howard J, Hyman AA (2003) Dynamics and mechanics of the microtubule plus end. Nature 422:753–758
29. Small JV, Kaverina I (2003) Microtubules meet substrate adhesions to arrange cell polarity. Curr Opin Cell Biol 15:40–47
30. Mitchison TJ (1993) Localization of an exchangeable GTP binding site at the plus end of microtubules. Science 261:1044–1047
31. Desai A, Mitchison TJ (1997) Microtubule polymerization dynamics. Annu Rev Cell Dev Biol 13:83–117
32. Nogales E, Whittaker M, Milligan RA, Downing KH (1999) High-resolution model of the microtubule. Cell 96:79–88
33. Ravelli RB, Gigant B, Curmi PA, Jourdain I, Lachkar S, Sobel A, Knossow M (2004) Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. Nature 428:198–202
34. Kumar MD, Gromiha MM (2006) PINT: protein-protein interactions thermodynamic database. Nucleic Acids Res 34:D195–D198
35. Tung M, Gallagher DT (2009) The biomolecular crystallization database version 4: expanded content and new features. Acta Crystallogr D Biol Crystallogr 65:18–23
36. Grigorieff N, Harrison SC (2011) Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. Curr Opin Struct Biol 21:265–273
37. Lupetti P (2005) Cryotechniques for electron microscopy: a mini review. In: From Cells to Protein: Imaging Nature across Dimensions, Springer, Dordrecht, pp 53–70
38. Zhang X, Settembre E, Xu C, Dormitzer PR, Bellamy R, Harrison SC, Grigorieff N (2008) Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. Proc Natl Acad Sci USA 105:1867–1872
39. Zhou ZH (2008) Towards atomic resolution structural determination by single-particle cryo-electron microscopy. Curr Opin Struct Biol 18:218–228

40. Jonic S, Venien-Bryan C (2009) Protein structure determination by electron cryo-microscopy. Curr Opin Pharmacol 9:636–642
41. Baker ML, Zhang J, Ludtke SJ, Chiu W (2010) Cryo-EM of macromolecular assemblies at near-atomic resolution. Nat Protoc 5:1697–1708
42. Zhou ZH (2011) Atomic resolution cryo electron microscopy of macromolecular complexes. Adv Protein Chem Struct Biol 82:1–35
43. Ball G, Parton RM, Hamilton RS, Davis I (2012) A cell biologist's guide to high resolution imaging. Methods Enzymol 504:29–55
44. Geerlof A, Brown J, Coutard B, Egloff MP, Enguita FJ, Fogg MJ, Gilbert RJ, Groves MR, Haouz A, Nettleship JE, Nordlund P, Owens RJ, Ruff M, Sainsbury S, Svergun DI, Wilmanns M (2006) The impact of protein characterization in structural proteomics. Acta Crystallogr D Biol Crystallogr 62:1125–1136
45. Banci L, Bertini I, Mangani S (2005) Integration of XAS and NMR techniques for the structure determination of metalloproteins: examples from the study of copper transport proteins. J Synchrotron Radiat 12:94–97
46. Strange RW, Feiters MC (2008) Biological X-ray absorption spectroscopy (BioXAS): a valuable tool for the study of trace elements in the life sciences. Curr Opin Struct Biol 18:609–616
47. Banci L, Bertini I, Del Conte R, Mangani S, Meyer-Klaucke W (2003) X-ray absorption and NMR spectroscopic studies of CopZ, a copper chaperone in Bacillus subtilis: the coordination properties of the copper ion. Biochemistry 42:2467–2474
48. Phelps RA, Cann JR (1957) Effect of binding of ions and other small molecules on protein structure. III. Influence of amino acids on the isomerization of proteins. J Am Chem Soc 79:4677–4679
49. Fegan A, White B, Carlson JCT, Wagner CR (2010) Chemically controlled protein assembly: techniques and applications. Chem Rev 110:3315–3336
50. Nooren IM, Thornton JM (2003) Diversity of protein–protein interactions. EMBO J 22:3486–3492
51. Cardinale D, Salo-Ahen OM, Ferrari S, Ponterini G, Cruciani G, Carosati E, Tochowicz AM, Mangani S, Wade RC, Costi MP (2010) Homodimeric enzymes as drug targets. Curr Med Chem 17:826–846
52. Xu X, Song Y, Li Y, Chang J, Zhang H, An L (2010) The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. Protein Expr Purif 72:149–156
53. Li Y (2011) The tandem affinity purification technology: an overview. Biotechnol Lett 33:1487–1499
54. Williamson MP, Sutcliffe MJ (2010) Protein-protein interactions. Biochem Soc Trans 38:875–878
55. Kerrigan JJ, Xie Q, Ames RS, Lu Q (2011) Production of protein complexes via co-expression. Protein Expr Purif 75:1–14
56. Zhao Y, Bishop B, Clay JE, Lu W, Jones M, Daenke S, Siebold C, Stuart DI, Yvonne Jones E, Radu Aricescu A (2011) Automation of large scale transient protein expression in mammalian cells. J Struct Biol 175:209–215
57. Vijayachandran LS, Viola C, Garzoni F, Trowitzsch S, Bieniossek C, Chaillet M, Schaffitzel C, Busso D, Romier C, Poterszman A, Richmond TJ, Berger I (2011) Robots, pipelines, polyproteins: enabling multiprotein expression in prokaryotic and eukaryotic cells. J Struct Biol 175:198–208
58. Perrakis A, Romier C (2008) Assembly of protein complexes by coexpression in prokaryotic and eukaryotic hosts: an overview. Methods Mol Biol 426:247–256
59. Busso D, Peleg Y, Heidebrecht T, Romier C, Jacobovitch Y, Dantes A, Salim L, Troesch E, Schuetz A, Heinemann U, Folkers GE, Geerlof A, Wilmanns M, Polewacz A, Quedenau C, Büssow K, Adamson R, Blagova E, Walton J, Cartwright JL, Bird LE, Owens RJ, Berrow NS, Wilson KS, Sussman JL, Perrakis A, Celie PHN (2011) Expression of protein

complexes using multiple Escherichia coli protein co-expression systems: a benchmarking study. J Struct Biol 175:159–170
60. Bieniossek C, Nie Y, Frey D, Olieric N, Schaffitzel C, Collinson I, Romier C, Berger P, Richmond TJ, Steinmetz MO, Berger I (2009) Automated unrestricted multigene recombineering for multiprotein complex production. Nat Methods 6:447–450
61. Romier C, Ben JM, Albeck S, Buchwald G, Busso D, Celie PH, Christodoulou E, De M, van GS, Knipscheer P, Lebbink JH, Notenboom V, Poterszman A, Rochel N, Cohen SX, Unger T, Sussman JL, Moras D, Sixma TK, Perrakis A (2006) Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. Acta Crystallogr D Biol Crystallogr 62:1232–1242
62. Passon DM, Lee M, Fox AH, Bond CS (2011) Crystallization of a paraspeckle protein PSPC1-NONO heterodimer. Acta Crystallogr Sect F 67:1231–1234
63. Passon DM, Lee M, Rackham O, Stanley WA, Sadowska A, Filipovska A, Fox AH, Bond CS (2012) Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. Proc Natl Acad Sci USA 109:4846–4850
64. Lutter R, Abrahams JP, van Raaij MJ, Todd RJ, Lundqvist T, Buchanan SK, Leslie AG, Walker JE (1993) Crystallization of F1-ATPase from bovine heart mitochondria. J Mol Biol 229:787–790
65. Leslie AG, Abrahams JP, Braig K, Lutter R, Menz RI, Orriss GL, van Raaij MJ, Walker JE (1999) The structure of bovine mitochondrial F1-ATPase: an example of rotary catalysis. Biochem Soc Trans 27:37–42
66. Gibbons C, Montgomery MG, Leslie AG, Walker JE (2000) The structure of the central stalk in bovine F(1)-ATPase at 2.4 Å resolution. Nat Struct Biol 7:1055–1061
67. Stock D, Leslie AG, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. Science 286:1700–1705
68. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, Nakashima R, Yaono R, Yoshikawa S (1996) The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. Science 272:1136–1144
69. Xia D, Yu CA, Kim H, Xia JZ, Kachurin AM, Zhang L, Yu L, Deisenhofer J (1997) Crystal structure of the cytochrome bc1 complex from bovine heart mitochondria. Science 277:60–66
70. Morth JP, Pedersen BP, Toustrup-Jensen MS, Sorensen TL, Petersen J, Andersen JP, Vilsen B, Nissen P (2007) Crystal structure of the sodium-potassium pump. Nature 450:1043–1049
71. Drew D, Froderberg L, Baars L, de Gier JW (2003) Assembly and overexpression of membrane proteins in Escherichia coli. Biochim Biophys Acta 1610:3–10
72. Miroux B, Walker JE (1996) Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260(19):289–298
73. Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, Högbom M, van Wijk KJ, Slotboom DJ, Persson JO, de Gier JW (2008) Tuning Escherichia coli for membrane protein overexpression. PNAS 105:14371–14376
74. Hays FA, Roe-Zurz Z, Li M, Kelly L, Gruswitz F, Sali A, Stroud RM (2009) Ratiocinative screen of eukaryotic integral membrane protein expression and solubilization for structure determination. J Struct Funct Genomics 10:9–16
75. Gu S, Rehman S, Wang X, Shevchik VE, Pickersgill RW (2012) Structural and functional insights into the pilotin-secretin complex of the type II secretion system. PLoS Pathog 8:e1002531
76. Romes EM, Tripathy A, Slep KC (2012) The Structure of a yeast Dyn2-Nup159 complex and the molecular basis for the dynein light chain: nuclear pore interaction. J Biol Chem 287(19):15862–15873
77. Su CC, Long F, Zimmermann MT, Rajashankar KR, Jernigan RL, Yu EW (2011) Crystal structure of the CusBA heavy-metal efflux complex of Escherichia coli. Nature 470:558–562

78. Su CC, Yang F, Long F, Reyon D, Routh MD, Kuo DW, Mokhtari AK, Van Ornam JD, Rabe KL, Hoy JA, Lee YJ, Rajashankar KR, Yu EW (2009) Crystal structure of the membrane fusion protein CusB from Escherichia coli. J Mol Biol 393:342–355
79. Long F, Su CC, Zimmermann MT, Boyken SE, Rajashankar KR, Jernigan RL, Yu EW (2010) Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. Nature 467:484–488
80. O'Halloran TV, Culotta VC (2000) Metallochaperones: an intracellular shuttle service for metal ions. J Biol Chem 275:25057–25060
81. Arnesano F, Banci L, Bertini I, Ciofi-Baffoni S, Molteni E, Huffman DL, O'Halloran TV (2002) Metallochaperones and metal-transporting ATPases: a comparative analysis of sequences and structures. Genome Res 12:255–271
82. Banci L, Bertini I, Cantini F, Felli IC, Gonnelli L, Hadjiliadis N, Pierattelli R, Rosato A, Voulgaris P (2006) The Atx1-Ccc2 complex is a metal-mediated protein–protein interaction. Nat Chem Biol 2:367–368
83. Bertini I, Cavallaro G, McGreevy KS (2010) Cellular copper management—a draft user's guide. Coord Chem Rev 254:506–524
84. Lamb AL, Torres AS, O'Halloran TV, Rosenzweig AC (2001) Heterodimeric structure of superoxide dismutase in complex with its metallochaperone. Nat Struct Biol 8:751–755
85. Banci L, Bertini I, Calderone V, la-Malva N, Felli IC, Neri S, Pavelkova A, Rosato A (2009) Copper(I)-mediated protein–protein interactions result from suboptimal interaction surfaces. Biochem J 422:37–42
86. Noinaj N, Easley NC, Oke M, Mizuno N, Gumbart J, Boura E, Steere AN, Zak O, Aisen P, Tajkhorshid E, Evans RW, Gorringe AR, Mason AB, Steven AC, Buchanan SK (2012) Structural basis for iron piracy by pathogenic Neisseria. Nature 483:53–58
87. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le TI, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: AG protein-coupled receptor. Science 289:739–745
88. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y (2002) Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. Proc Natl Acad Sci USA 99:5982–5987
89. Massotte D (2003) G protein-coupled receptor overexpression with the baculovirus—insect cell system: a tool for structural and functional studies. Biochimica et Biophysica Acta (BBA): Biomembr 1610:77–89
90. Lanyi JK, Schobert B (2004) Local-global conformational coupling in a heptahelical membrane protein: transport mechanism from crystal structures of the nine states in the bacteriorhodopsin photocycle. Biochemistry 43:3–8
91. Lundstrom K (2005) Structural biology of G protein-coupled receptors. Bioorg Med Chem Lett 15:3654–3657
92. Lundstrom K (2005) Structural genomics of GPCRs. Trends Biotechnol 23:103–108
93. Rasmussen SGF, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF, Schertler GFX, Weis WI, Kobilka BK (2007) Crystal structure of the human [bgr]2 adrenergic G-protein-coupled receptor. Nature 450:383–387
94. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318:1258–1265
95. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. Nature 454:486–491
96. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin. Nature 454:183–187
97. Warne T, Serrano-Vega MJ, Tate CG, Schertler GF (2009) Development and crystallization of a minimal thermostabilised G protein-coupled receptor. Protein Expr Purif 65:204–213

98. Martin A, Damian M, Laguerre M, Parello J, Pucci B, Serre L, Mary S, Marie J, Baneres JL (2009) Engineering a G protein-coupled receptor for structural studies: stabilization of the BLT1 receptor ground state. Protein Sci 18:727–734
99. Kimple AJ, Soundararajan M, Hutsell SQ, Roos AK, Urban DJ, Setola V, Temple BR, Roth BL, Knapp S, Willard FS, Siderovski DP (2009) Structural determinants of G-protein alpha subunit selectivity by regulator of G-protein signaling 2 (RGS2). J Biol Chem 284:19402–19411
100. Bokoch MP, Zou Y, Rasmussen SG, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. Nature 463:108–112
101. Wacker D, Fenalti G, Brown MA, Katritch V, Abagyan R, Cherezov V, Stevens RC (2010) Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. J Am Chem Soc 132:11443–11445
102. Moukhametzianov R, Warne T, Edwards PC, Serrano-Vega MJ, Leslie AG, Tate CG, Schertler GF (2011) Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. Proc Natl Acad Sci USA 108:8228–8232
103. Zou Y, Weis WI, Kobilka BK (2012) N-Terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. PLoS ONE 7:e46039
104. Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, Reinhart G, Desale H, Clemons B, Cahalan SM, Schuerer SC, Sanna MG, Han GW, Kuhn P, Rosen H, Stevens RC (2012) Crystal structure of a lipid G protein-coupled receptor. Science 335:851–855
105. Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, Nakada-Nakura Y, Kusano-Arai O, Weyand S, Shimamura T, Nomura N, Cameron AD, Kobayashi T, Hamakubo T, Iwata S, Murata T (2012) G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. Nature 482:237–240
106. Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK (2011) Crystal structure of the beta2 adrenergic receptor-Gs protein complex. Nature 477(19):549–555
107. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK (2007) GPCR engineering yields high-resolution structural insights into beta-2-Adrenergic receptor function. Science 318:1266–1273
108. Hunte C, Zickermann V, Brandt U (2010) Functional modules and structural basis of conformational coupling in mitochondrial complex I. Science 329:448–451
109. Sun F, Huo X, Zhai Y, Wang A, Xu J, Su D, Bartlam M, Rao Z (2005) Crystal structure of mitochondrial respiratory membrane protein complex II. Cell 121:1043–1057
110. Lancaster CR, Kroger A, Auer M, Michel H (1999) Structure of fumarate reductase from Wolinella succinogenes at 2.2 Å resolution. Nature 402:377–385
111. Iverson TM, Luna-Chavez C, Cecchini G, Rees DC (1999) Structure of the Escherichia coli fumarate reductase respiratory complex. Science 284:1961–1966
112. Lange C, Nett JH, Trumpower BL, Hunte C (2001) Specific roles of protein-phospholipid interactions in the yeast cytochrome bc1 complex structure. EMBO J 20:6591–6600
113. Menz RI, Walker JE, Leslie AG (2001) Structure of bovine mitochondrial F(1)-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. Cell 106:331–341
114. Wimberly BT, Brodersen DE, Clemons WM Jr, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V (2000) Structure of the 30S ribosomal subunit. Nature 407:327–339
115. Rabl J, Leibundgut M, Ataide SF, Haag A, Ban N (2011) Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. Science 331:730–736
116. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 289:905–920

117. Klinge S, Voigts-Hoffmann F, Leibundgut M, Arpagaus S, Ban N (2011) Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. Science 334:941–948
118. Groll M, Ditzel L, Lowe J, Stock D, Bochtler M, Bartunik HD, Huber R (1997) Structure of 20S proteasome from yeast at 2.4 Å resolution. Nature 386:463–471
119. Chayen NE (2004) Turning protein crystallisation from an art into a science. Curr Opin Struct Biol 14:577–583
120. Chayen NE, Saridakis E (2008) Protein crystallization: from purified protein to diffraction-quality crystal. Nat Methods 5:147–153
121. Bolanos-Garcia VM, Chayen NE (2009) New directions in conventional methods of protein crystallization. Prog Biophys Mol Biol 101:3–12
122. Nooren IM, Thornton JM (2003) Structural characterisation and functional significance of transient protein–protein interactions. J Mol Biol 325:991–1018
123. Russell RB, Alber F, Aloy P, Davis FP, Korkin D, Pichaud M, Topf M, Sali A (2004) A structural perspective on protein–protein interactions. Curr Opin Struct Biol 14:313–324
124. Radaev S, Li S, Sun PD (2006) A survey of protein–protein complex crystallizations. Acta Crystallogr D Biol Crystallogr 62:605–612
125. Radaev S, Sun PD (2002) Crystallization of protein–protein complexes. J Appl Crystallogr 35:674–676
126. Toogood PL (2002) Inhibition of protein–protein association by small molecules: approaches and progress. J Med Chem 45:1543–1558
127. Pagliaro L, Felding J, Audouze K, Nielsen SJ, Terry RB, Krog-Jensen C, Butcher S (2004) Emerging classes of protein–protein interaction inhibitors and new tools for their development. Curr Opin Chem Biol 8:442–449
128. Bourgeas R, Basse MJ, Morelli X, Roche P (2010) Atomic analysis of protein–protein interfaces with known inhibitors: the 2P2I database. PLoS ONE 5:e9598
129. Morelli X, Bourgeas R, Roche P (2011) Chemical and structural lessons from recent successes in protein–protein interaction inhibition (2P2I). Curr Opin Chem Biol 15:475–481
130. Sattler M, Liang H, Nettesheim D, Meadows RP, Harlan JE, Eberstadt M, Yoon HS, Shuker SB, Chang BS, Minn AJ, Thompson CB, Fesik SW (1997) Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. Science 275:983–986
131. Liu Z, Sun C, Olejniczak ET, Meadows RP, Betz SF, Oost T, Herrmann J, Wu JC, Fesik SW (2000) Structural basis for binding of Smac/DIABLO to the XIAP BIR3 domain. Nature 408:1004–1008
132. Agamennone M, Cesari L, Lalli D, Turlizzi E, Del CR, Turano P, Mangani S, Padova A (2010) Fragmenting the S100B–p53 interaction: combined virtual/biophysical screening approaches to identify ligands. ChemMedChem 5:428–435
133. Cossu F, Milani M, Mastrangelo E, Vachette P, Servida F, Lecis D, Canevari G, Delia D, Drago C, Rizzo V, Manzoni L, Seneci P, Scolastico C, Bolognesi M (2009) Structural basis for bivalent Smac-mimetics recognition in the IAP protein family. J Mol Biol 392:630–644
134. Lee EF, Czabotar PE, Smith BJ, Deshayes K, Zobel K, Colman PM, Fairlie WD (2007) Crystal structure of ABT-737 complexed with Bcl-xL: implications for selectivity of antagonists of the Bcl-2 family. Cell Death Differ 14:1711–1713
135. He MM, Smith AS, Oslob JD, Flanagan WM, Braisted AC, Whitty A, Cancilla MT, Wang J, Lugovskoy AA, Yoburn JC, Fung AD, Farrington G, Eldredge JK, Day ES, Cruz LA, Cachero TG, Miller SK, Friedman JE, Choong IC, Cunningham BC (2005) Small-molecule inhibition of TNF-alpha. Science 310:1022–1025
136. Silvian LF, Friedman JE, Strauch K, Cachero TG, Day ES, Qian F, Cunningham B, Fung A, Sun L, Shipps GW, Su L, Zheng Z, Kumaravel G, Whitty A (2011) Small molecule inhibition of the TNF family cytokine CD40 ligand through a subunit fracture mechanism. ACS Chem Biol 6:636–647
137. Cardinale D, Guaitoli G, Tondi D, Luciani R, Henrich S, Salo-Ahen OM, Ferrari S, Marverti G, Guerrieri D, Ligabue A, Frassineti C, Pozzi C, Mangani S, Fessas D, Guerrini R, Ponterini G, Wade RC, Costi MP (2011) Protein-protein interface-binding peptides inhibit

the cancer therapy target human thymidylate synthase. Proc Natl Acad Sci USA 108:E542–E549
138. Thangudu RR, Tyagi M, Shoemaker BA, Bryant SH, Panchenko AR, Madej T (2010) Knowledge-based annotation of small molecule binding sites in proteins. BMC Bioinform 11:365
139. Shoemaker BA, Zhang D, Tyagi M, Thangudu RR, Fong JH, Marchler-Bauer A, Bryant SH, Madej T, Panchenko AR (2012) IBIS (inferred biomolecular interaction server) reports, predicts and integrates multiple types of conserved interactions for proteins. Nucleic Acids Res 40:D834–D840
140. Dynek JN, Vucic D (2010) Antagonists of IAP proteins as cancer therapeutics. Cancer Lett 332(2):206–214
141. Fulda S, Vucic D (2012) Targeting IAP proteins for therapeutic intervention in cancer. Nat Rev Drug Discov 11:109–124
142. Wilson WH, O'Connor OA, Czuczman MS, LaCasce AS, Gerecitano JF, Leonard JP, Tulpule A, Dunleavy K, Xiong H, Chiu YL, Cui Y, Busman T, Elmore SW, Rosenberg SH, Krivoshik AP, Enschede SH, Humerickhouse RA (2010) Navitoclax, a targeted high-affinity inhibitor of BCL-2, in lymphoid malignancies: a phase 1 dose-escalation study of safety, pharmacokinetics, pharmacodynamics, and antitumour activity. Lancet Oncol 11:1149–1159
143. Roberts AW, Seymour JF, Brown JR, Wierda WG, Kipps TJ, Khaw SL, Carney DA, He SZ, Huang DC, Xiong H, Cui Y, Busman TA, McKeegan EM, Krivoshik AP, Enschede SH, Humerickhouse R (2012) Substantial susceptibility of chronic lymphocytic leukemia to BCL2 inhibition: results of a phase I study of navitoclax in patients with relapsed or refractory disease. J Clin Oncol 30:488–496
144. Rudin CM, Hann CL, Garon EB, de Ribeiro OM, Bonomi PD, Camidge DR, Chu Q, Giaccone G, Khaira D, Ramalingam SS, Ranson MR, Dive C, McKeegan EM, Chyla BJ, Dowell BL, Chakravartty A, Nolan CE, Rudersdorf N, Busman TA, Mabry MH, Krivoshik AP, Humerickhouse RA, Shapiro GI, Gandhi L (2012) Phase II study of single-agent navitoclax (ABT-263) and biomarker correlates in patients with relapsed small cell lung cancer. Clin Cancer Res 18:3163–3169
145. Weber L (2010) Patented inhibitors of p53-Mdm2 interaction (2006–2008). Expert Opin Ther Pat 20:179–191
146. Millard M, Pathania D, Grande F, Xu S, Neamati N (2011) Small-molecule inhibitors of p53-MDM2 interaction: the 2006–2010 update. Curr Pharm Des 17:536–559
147. Warner WA, Sanchez R, Dawoodian A, Li E, Momand J (2012) Identification of FDA-approved drugs that computationally Bind to MDM2. Chem Biol Drug Des 80:631–637

Chapter 6
Fluorescence Observables and Enzyme Kinetics in the Investigation of PPI Modulation by Small Molecules: Detection, Mechanistic Insight, and Functional Consequences

Glauco Ponterini

6.1 Introduction

Protein homo- or hetero-oligomers are widespread in living systems. In many instances, their functional roles have been established or conjectured [1, 2], and the potential impact of their study on the understanding of apparently unrelated phenomena, such as protein folding and inter-domain interaction, has been underlined [3]. The discovery of small molecules able to modulate protein–protein interactions (PPIs), hence changing the stability of the oligomers, may therefore assume a strong functional significance. However, PPIs often involve extended protein interfaces which have long been perceived as chemically featureless; therefore, designing small molecules with a potential to interfere with high affinity and specificity with a multiprotein complex, and modulate its stability, may represent a challenging chemical problem having functional biological implications [2, 4–7]. Fortunately, the above perception is increasingly refuted [8, 9], and examples of small compounds able to bind to a protein surface with affinities high enough to compete with the binding of the protein to other proteins are rapidly accumulating [2, 4, 6–8, 10 and Chap. 2 in this book]. A qualitative/quantitative characterization of the effects of these compounds on protein–protein complex formation is essential in drug discovery processes targeting PPIs and, as we shall see, has been obtained through several experimental approaches. On the other hand, addressing and characterizing the mechanism of action of such small molecules, including the relevant structural features, often remains a prohibitive task. However, it is this higher level of knowledge that, while answering intrinsically relevant questions related, for example, with the types and number of PPIs involved, may provide valuable hints

G. Ponterini (✉)
Department of Life Sciences, University of Modena and Reggio Emilia,
via Campi 183, 41125 Modena, Italy
e-mail: glauco.ponterini@unimore.it

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4_6,

for designing new small molecules characterized by stronger, or more specific, modulating actions [10].

Essentially, any biophysical experimental observable, whether associated with light-matter or electron–matter interactions (absorption, both molecular and plasmonic, emission, scattering, diffraction), with heat exchange, with molecular-size-dependent radial distribution in a rotating centrifuge, or migration in a chromatographic column or under an electric field, can in principle be employed to investigate PPIs, and many of them have actually been so [11–14]. On the other hand, most reported examples of modulation of these interactions by small molecules are based on experimental methods that probe a narrower group of biophysical observables, especially those adaptable for protocols with high-throughput-screening capacities [15]. Designing and setting up an experimental method that is able not only to give a signal when testing an active compound, but also to provide structural and/or functional insight on the effects of its binding on PPIs, that is, on the structure/stability of supramolecular protein assemblies, can be severely challenging. Such a method must fulfill the requirements for detecting and characterizing PPIs, but with a few important additions: It should be sensitive enough to reflect changes in these interactions that (1) may be intrinsically small and (2) are often produced in ligand/protein complexes present at much lower concentrations than the unperturbed protein oligomers. This requirement is particularly severe in the many examples of functional relevance in which transient multimeric protein assemblies are addressed and weak to moderate PPIs that govern their formation are modulated [16]. Finally, and rarely achieved, (3) the experimental approach should also provide mechanistic insight, that is, it should highlight the changes in structure and behavior brought about in the protein assembly by the binding to the small molecule.

In the reported examples of PPI modulation, spectroscopic methods are most often employed for the purpose, being naturally connected with the structure of the molecules probed and the interactions they experience and, at least in many favorable cases, because of their good sensitivity and specificity. Noteworthy exceptions are represented by studies based on size-exclusion chromatography (SEC), sedimentation equilibrium analytical ultracentrifugation, usually in the sedimentation velocity mode of measurement [17], and surface plasmon resonance (SPR). The first two techniques are based on a spatial separation of protein monomers and different multimers and have found some applications in the field of PPI modulation screening (some cases are described in Ref. [18]). However, because of complicating effects (e.g., in SEC, nonspecific binding of differently aggregated proteins to the stationary phase may differentially affect the chemical potentials of monomers and different multimers [19]), and, more generally, because of the competing tendency of the system to locally re-establish the aggregation equilibrium, thus blurring the desired size-based spatial separation, these methods require an optimization of the experimental conditions and a careful calibration and data analysis to yield reliable quantitative data [20, 21]. SPR has become a widespread method to monitor formation/dissociation of protein complexes [22] and, in a few examples, to screen the effects thereon of libraries of

small molecules (see, e.g., Refs. [23] and [24]). In this technique, the rate constants for association and dissociation of an added compound to an immobilized partner are usually evaluated from the time evolution of the observable, and, from their ratios, binding equilibrium constants are estimated [25]. SPR suffers from a few limiting features; for example, irreversibility of ligand/target complex formation and, relevant to our subject, the occurrence of other processes following binding, such as protein–protein complex formation or dissociation, represent undesired events that complicate data fitting. Also, the SPR observable consists in shifts in the resonance wavelength of gold surface plasmons caused by the binding of organic molecules on the metal surface. It is not, therefore, a molecular spectroscopic technique and can hardly provide a molecular-scale insight comparable with that obtainable from spectroscopic tools. Among the latter, methods based on nuclear magnetic resonance (NMR) [16, 26] and on molecular fluorescence (see, e.g., Ref. [27]) are the most powerful and most often employed. Both families of approaches can provide direct insight into the structural details and the dynamics of protein–protein and protein–ligand complexes, though with their specificities, advantages, and limitations. NMR applications are described elsewhere in this volume (Rebecca Del Conte, Daniela Lalli, Paola Turano, NMR as a tool to target PPIs). Here, we have collected and commented on some representative examples of the potential of fluorescence-based methods in the screening and molecular-scale mechanistic investigation of PPI modulation by small molecules (this overview is intended to be by way of example rather than exhaustive).

From a quite different point of view, when multimeric enzymes are involved, kinetic analyses can be employed to screen the functional consequences of the modulation of PPIs by small molecules and, more relevant in the perspective of this contribution, to test mechanistic hypotheses. The potential of this experimental opportunity has probably been overlooked. In the final part of this chapter, we will briefly and critically review some relevant examples.

6.2 Fluorescence Observables

These experimental approaches take advantage of a variety of observables. Properties such as spectra, intensities (related to quantum yields), time-decays, and anisotropies of intrinsic protein fluorophores, of extrinsic fluorescent tags and, even, of the same small molecules added to modulate PPIs, as well as phenomena such as static and collisional quenching, including electron and excitation-energy transfer, or exciton interaction, whose efficiencies crucially depend on the distance between the partners and their relative orientation [28, 29], may in principle be used to monitor changes in the protein aggregation pattern.

6.2.1 Protein Fluorescence

Changes in properties of the intrinsic fluorescence of proteins related with changes in their aggregation state (Fig. 6.1) have been reported, and exploited, in quite a few instances. Pertinent examples are the cases of interferon-γ dimer/monomer transition, of calmodulin interaction with a neuronal target protein (see Ref. [29], Chap. 16), of melittin self-association (see Ref. [29], Chap. 17), and of the complexation of a retinal phosphodiesterase subunit with two subunits of heterotrimeric G-protein transducin [30].

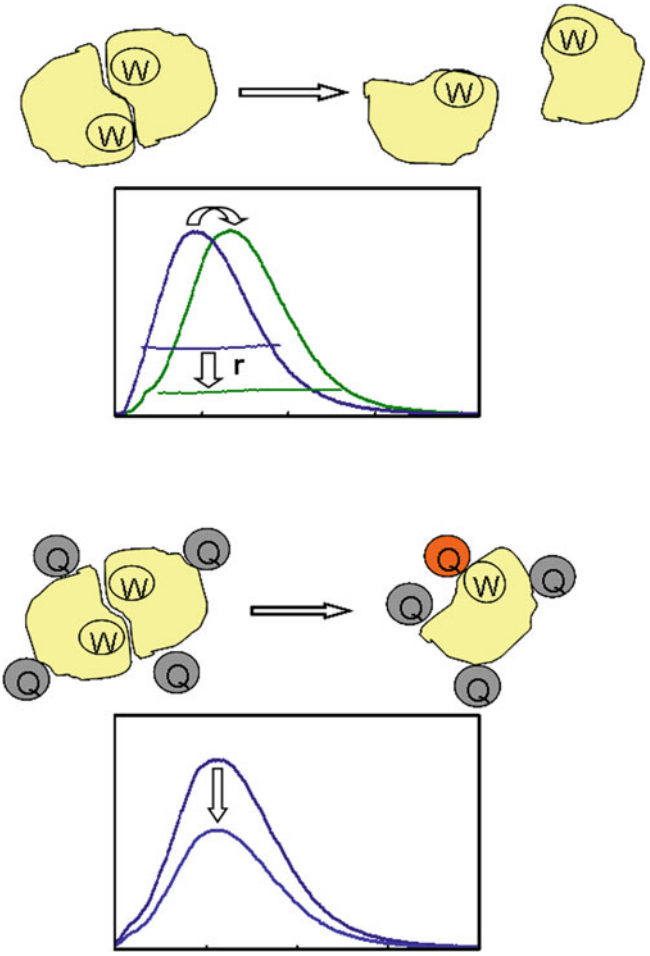


Fig. 6.1 Protein multimer disruption and steady-state intrinsic (tryptophan, *W*) protein fluorescence. *Top* emission spectral shift associated with a change in *W* environment; change in anisotropy related with a change in rotational mobility. *Bottom* decrease in emission intensity due to increased accessibility of external quenchers (*Q*) to *W*s

Because of the different electronic distributions of the lowest excited (S_1) and ground electronic states, of the possible involvement of $n\pi^*$ states and of possible free-volume-requiring S_1-state relaxation processes, emission spectra are often sensitive to the polarity, proticity, and microviscosity of the fluorophore environment. The main intrinsic fluorophore of proteins, tryptophan, exhibits such a sensitivity: Its emission shows a bathochromic shift with increasing polarity of the local environment (upper panel in Fig. 6.1), and a blurring of the vibronic structure when moving from an aprotic to a protic environment, associated with the stabilization of the ^1L_a relative to the ^1L_b states (see Ref. [29], Chap. 16). Thus, the intrinsic protein emission spectrum can be used to monitor changes in the solvent exposure of the tryptophan residues. In the only example known to us of PPI modulation by small molecules monitored in this way, shifts of the emission maximum of glutamate dehydrogenase enabled a reversible hexamer-to-trimer dissociation of the protein to be observed. The approach, based on dynamic light scattering, aimed at progressively disrupting PPIs using guanidinium hydrochloride at low concentrations. Increases in ΔGs were estimated for the process upon binding of norvaline and glutamate to the protein, indicating a ligand-induced stabilization of the hexamer [31], a rare example of positive modulation of PPIs by small molecules.

Measuring changes in intensity of the intrinsic protein fluorescence in complexes relative to the corresponding separated components seems quite a simple way to monitor complex formation or disruption (see, e.g., Ref. [30]). In spite of this, we are aware of only one example of small-molecule-induced changes in protein complexation investigated this way. A combination of SEC and intrinsic protein fluorescence measurement showed that tethered peptides, corresponding to the N- and C-termini of HIV-1 protease, targeted the dimer interface of HIV-1 protease and decreased the fraction of enzyme dimer in solution [32]. Here, the presence of a tryptophan near the monomer/monomer interface was exploited: Addition of the tethered dipeptide inhibitor to the protein caused a marked fluorescence quenching that was not observed with a conventional active-site inhibitor and was presumed to be due to an 'increased solvent exposure' of this tryptophan in the monomers. This statement is probably misleading, as it suggests that exposed tryptophans are more likely quenched than are more buried ones. This is true when accessibility by external quenchers, such as acrylamide or oxygen, is concerned (see Ref. [29], Chap. 16 and the following lines). However, it is well known that the lifetimes and quantum yields of tryptophans in proteins are controlled by a number of quenching processes that involve several different residues, as well as peptide bonds of the backbone [33]. As a result, lifetimes and quantum yields do not correlate with emission maxima, that is, with the solvent exposure of the tryptophans (see Ref. [29], Chap. 16). So, in principle, these observables could be exploited to monitor changes in the aggregation state of the proteins if these result in structural changes that occur in the proximity of tryptophan residues, even buried ones, and that, in turn, affect the efficiency of the quenching processes.

Protein fluorescence quantum yields are reduced in the presence of dissolved quenchers that can access one or more tryptophans. Thus, measurements of protein

emission intensity in the presence of dissolved quenchers can reveal changes in accessibility resulting from formation or disruption of protein complexes (lower panel in Fig. 6.1). For example, increased quenching by KI of the intrinsic fluorescence of *Plasmodium falciparum* triosephosphate isomerase, following mutation of a tyrosine at the subunit interface to glycine, indicated a larger accessibility by the iodide quencher in water of a tryptophan residue near the interface, associated with dimer disruption, as confirmed by gel filtration experiments [34]. In a slightly different approach, the tetramer-to-dimer and dimer-to-monomer dissociation kinetics of three apolipoproteins of the E family were followed by observing the decrease in intensity of the intrinsic protein fluorescence following dilution of the proteins in a solution of acrylamide, a classical tryptophan quencher [35]. In spite of the potential of this fluorescence observable, we are not aware of the use of experiments based on differential accessibility to quenchers to test PPI perturbation by small molecules.

In general, the intrinsic steady-state protein fluorescence properties, most notably anisotropy (upper panel in Fig. 6.1), are little employed to monitor changes in the protein aggregation pattern caused by interaction with small molecules, in spite of the simplicity of these measurements. The small number of examples of this kind might imply that the tryptophan emission properties are rarely significantly affected by changes in the tertiary and quaternary structure of proteins. We are not, however, aware that this has ever been actually observed and explicitly reported.

6.2.2 Fluorescence of Probes

When fluorescent labels are employed, the source of information about changes in PPIs, or in protein assemblies, is a change in the probe fluorescence properties. In order to provide information of mechanistic relevance, such changes must be traceable to varied probe environment, accessibility to quenchers, proximity to other fluorophores or rotational mobility.

As an example of a change in a probe environment, a fluorescence assay has been designed to test the binding of a library of tetrapeptides, modeled on the N-terminus of the pro-apoptotic protein Smac, to the surface pocket of the BIR3 binding region of the anti-apoptotic XIAP protein. Here, a solvent-sensitive fluorogenic naphthalene-based dye was attached to a tetrapeptide through a thiol linkage and, upon binding to XIAP, underwent a solvatochromic emission shift and a change in emission intensity (upper panel in Fig. 6.2). These changes, or, more precisely, their reversal (lower panel in Fig. 6.2), were employed to monitor the displacement of the bound tagged peptide by other untagged tetrapeptides, and quantify the corresponding binding equilibrium constants [36].

Anisotropy changes, which reflect changes in rotational mobility of the fluorophore in the free and bound states or, with lower sensitivity, when bound to a protein in different aggregation states, have been employed to characterize small

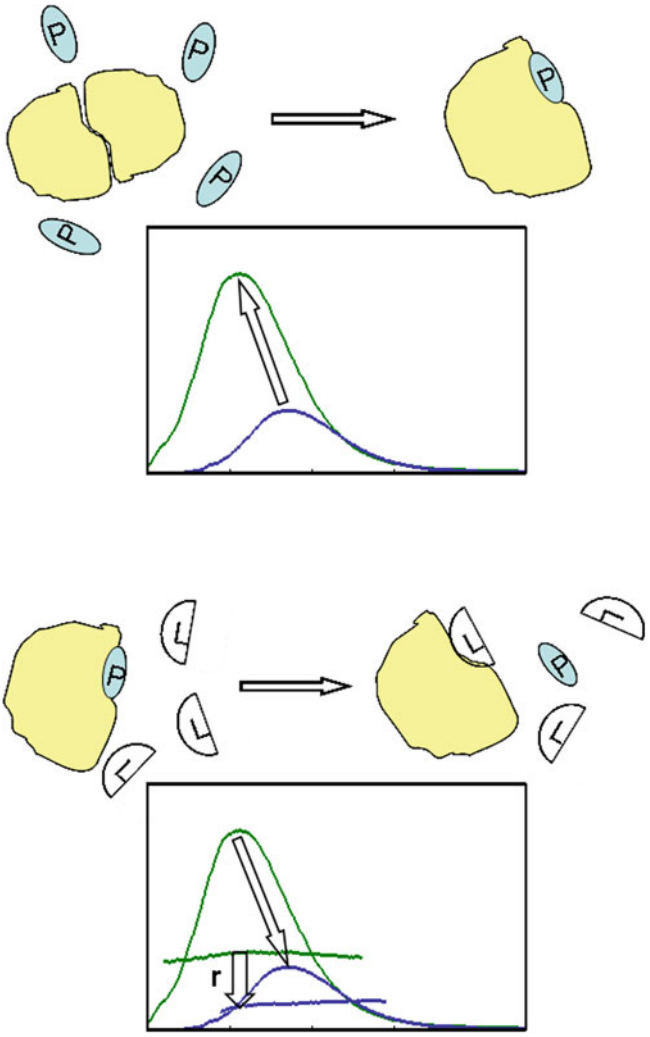


Fig. 6.2 *Top* protein multimer disruption and steady-state probe fluorescence: emission spectral shift and intensity change associated with a change in probe environment. *Bottom* reversed spectral changes and decrease in probe emission anisotropy caused by mass-law-governed displacement of the probe by a tested ligand (*L*). Here, *P* represents a fluorescent probe or, more often, a tagged peptide with good affinity for the protein–protein binding site

inhibitors of PPIs. An example of the latter kind is provided by self-association of a fluorescein-labeled retinoid-X-receptor to form tetramers that was followed by measuring the fluorescence anisotropy of the probe with increasing protein concentration. In the presence of 9-cis-retinoic acid, the final anisotropy was much lower than in its absence, an indication that formation of the tetramer, characterized by a slower rotational diffusion, was inhibited by this ligand [37]. A similar

fluorescence polarization assay, developed in a high-throughput format, was employed to screen compounds able to perturb the interaction between two peptides, designed from the binding regions of fibronectin and tissue transglutaminase, two proteins whose complex is believed to promote tumor cell adhesion and to be involved in the process of tumor dissemination [38]. One of the two peptides was tagged with a fluorescein molecule; addition of the other peptide resulted in a saturating increase in the anisotropy of the probe, due to complex formation. Small compounds able to inhibit formation of the complex caused a decrease in the observed anisotropy at fixed peptide concentrations.

The most widely employed fluorescence polarization assays are, however, based on competitive displacement of a labeled small molecule, often a peptide, known to bind at a region crucial for PPI (lower panel in Fig. 6.2). Rather than directly monitoring PPI modulation, the assay aims at testing the ability of small molecules to replace the labeled small molecule, which is assumed to mimic the partner protein. For example, some peptides designed from pro-apoptotic Smac were shown to bind to the BIR2 and BIR3 domains of the anti-apoptotic XIAP inhibitor protein [39–41]. This binding was quantitatively characterized by measuring the anisotropy of the emission from a peptide labeled with a carboxyfluorescein, which was progressively displaced from the BIR2 and BIR3 domains of XIAP by the tested unlabeled peptides in dose–response experiments. Similarly, the IC_{50}s of two small peptide-based inhibitors of the interaction between the von Hippel–Lindau protein (VHL), the substrate recognition subunit of an ubiquitin ligase, and its primary substrate, the hypoxia-inducible factor 1α (HIF-1α), were determined by measuring the anisotropy of a fluorescein-labeled HIF-1α peptide that binds VHL with a 560 nM affinity [42]. In another example, inhibition constants of several green tea polyphenols versus two Bcl-2 family anti-apoptotic proteins were determined with a competition assay based on dose–dependent displacement of a fluorescein-labeled peptide, reproducing the BH3 domain of the pro-apoptotic counterparts, and measurement of the resulting decrease in emission polarization of the fluorescein probe [43]. Essentially, the same approach has been employed [44, 45], also in a high-throughput version [46], to find inhibitors of the BH3/Bcl-2 interaction, to identify chelerythrine as an inhibitor of the Bcl-XL/BH3 complexation [47], to screen a series of terephthalamides as inhibitors of the Bcl-XL/BAK peptide binding [48], to test molecules, selected using a shape-comparison program, for activity against the ZipA–FtsZ interaction, an antibacterial target [49], and to identify a small inhibitor of the interaction between one of the proteins of the 14-3-3 family, implied in physiological and pathophysiological interactions with more than 200 proteins, and the pS259-Raf-1 peptide [50]. The additivity of fluorescence anisotropy [28] was crucial in some of these applications to enable the fractions of labeled peptides, bound and free, and the binding equilibrium constants to be determined. A problem with this kind of assay is that, in order to characterize quantitatively the binding of potent inhibitors, high-affinity labeled peptides, to be replaced by the tested compounds, must sometimes be designed and obtained for the purpose [41]. As a general comment on the widespread methods based on polarization of probe fluorescence, while

these experiments are efficient as bases for medium- to high-throughput screenings (HTS), because they rely on rotational mobility, which is controlled mainly by size and, to a lesser extent, shape, they generally lack structural/mechanistic insight on the protein/ligand binding modes.

Under favorable conditions, H-type exciton coupling between closely associated fluorescent labels may result in a marked emission quenching [51]. This phenomenon was exploited to monitor the dissociation kinetics of a subunit from trimeric tumor necrosis factor α induced by a small inhibitor of the protein and to deduce conclusions on the mechanism of the process [52]. Dissociation of a subunit caused loss of H-type exciton coupling between fluorescein molecules labeling different subunits and resulted in emission recovery.

A couple of examples of PPI modulation by small molecules are characterized by a hazy description of the molecular bases of the assay employed. In the first one, a change in fluorescence from a probe was only assumed to reflect changes in a protein assembly, but was not interpreted on a molecular level. The fluorescence of dansylated *L. casei* thymidylate synthase (TS) was monitored to investigate structural changes of the dimeric protein upon interaction with a 20-mer peptide designed to reproduce a sequence at the subunit interface [53]. Addition of this peptide, that inhibits TS, was found to result in a decrease in the emission intensity of a dansyl probe specifically bound to a Cys residue which resides at the dimer interface. However, both this quenching and the protein inhibition were attributed to a peptide-induced decrease in spectroscopically and kinetically observable labeled protein in solution due to aggregation/precipitation, rather than to interference of the peptide with protein dimer structure or stability. A similar lack of molecular-scale insight characterizes an affinity-based assay proposed as a screening tool for PPIs [54]. A change in fluorescence intensity of a 'generic probe' upon thermal denaturation of a protein to which it is bound was employed in a high-throughput miniaturized test. An increase in thermal stability was expected and observed as a consequence of the binding of a tested compound to the labeled target protein. This small-molecule binding possibly but not necessarily inhibited binding with other proteins.

A rare example of the use of fluorescence changes to investigate the ability of small molecules to inhibit protein–protein binding is provided by the competition between antimycin A and the pro-apoptotic proteins BAK, BAX, and BIK for binding to the hydrophobic grooves of Bcl-XL [55] or of a recombinant Bcl-2 [56], anti-apoptotic proteins overexpressed in many cancer cells. In these cases, it was an emission enhancement of antimycin A itself, that is, the small inhibitor of protein–protein association, that was employed to demonstrate its binding to the Bcl proteins. The assumption was that the emission quantum yield of this fluorophore is larger in the hydrophobic environment provided by the proteins. Conversely, in competition experiments, a decrease in emission from antimycin A was used to monitor binding of a nonfluorescent methoxy derivative to the same groove. In the second paper, parallel experiments were made with 1-anilino-8-naphthalene sulfonate, a widely employed hydrophobic probe with emission properties that are strongly environment dependent.

Another example of fluorescent small PPI modulator is a fluorene-based compound able to block the interaction between $\alpha 2\beta 1$ integrin and collagen, an interaction that has been shown to have an important role in thrombus formation and cancer spread [57]. The peculiarity and interest of this example come from the fact that this inhibitor was specifically designed in order not only to bind to the flat collagen-binding domain of the integrin but, also, to be fluorescent. This condition is not, however, sufficient to make a good self-probing small inhibitor. In addition, some property of the fluorescence must change upon binding of the small molecule to the target protein domain. In this case, a strong emission enhancement accompanied the binding, a result not easily predictable. A molecular structure composed of one or more fluorophores connected to a biologically active group through single bonds, which allow for some torsion, is probably a useful structural feature for a fluorophore that is desired to have its emission quantum yield increased in a constrained environment. A similar feature characterizes some well-known fluorescent DNA dyes, for example, Hoechst 33258 [58]. Finally, but importantly, some evidence must be available about the ability of the bound fluorescent molecule to modulate the binding affinity of the protein toward other proteins.

6.2.3 Bimolecular Processes: FRET

Measurement of the efficiency of fluorescence (or Förster) resonance energy transfer (FRET), a long known powerful method for obtaining information about molecular-scale distances [28], is the most widely employed fluorescence-based method to investigate protein–protein complexation equilibria. Typically, selective excitation of an excitation-energy donor results in emission from an acceptor with an efficiency that depends on distance and relative orientation. Measurement of this efficiency for a well-characterized donor–acceptor pair enables distance between the partners to be estimated under reasonable assumptions on their relative orientation. FRET between fluorescent partners, including proteins, can be quantitatively assessed both in steady-state and in time-resolved (TR) experiments [28, 59]. In the former, the donor and acceptor emissions are measured under continuous excitation, either as full spectra or, in higher-throughput screenings, as intensities at selected excitation and emission wavelengths. Steady-state FRET experiments may be employed to investigate the association/dissociation kinetics of multimeric proteins when these kinetics are slow relative to mixing/dilution times [35]. TR experiments consist in acquiring the fluorescence time profiles following pulsed excitation and analyzing them to derive FRET efficiencies from changes in fluorescence intensity decays (donors) or rises (acceptors). FRET experiments can also be performed on living cells by combining steady-state or TR fluorescence measurements with the spatial resolution of a conventional optical or a confocal microscope [28, 60]. As one of the many examples of experiments designed to monitor PPI in cells, steady-state FRET between two different mutants

of GFP fused to two human Four-and-a-half LIM-only proteins, FHL2 and FHL3, was employed to determine their interaction and to locate the site of this interaction in a single intact mammalian cell [61].

There follows a selection of examples of FRET-based experiments, designed to monitor modulation of PPIs by small molecules.

FRET from fluorescein to tetramethylrhodamine, selectively bound to two cysteine residues, each on a different monomer of dimeric human TS (Fig. 6.3), was exploited in concentration-dependent steady-state fluorescence measurements to determine the fraction of dimeric protein at each total protein concentration and, as a result, the monomer/dimer equilibrium constant [62]. The assay was then used to test whether some octapeptides, found to inhibit hTS through an unconventional mechanism, were able to disrupt the protein dimer [63]. The experimental results showed only a minor perturbation of FRET, consistent with the crystallographic evidence of a binding of the peptides at the subunit interface without causing significant destabilization of the protein dimer.

A small library obtained by computational interrogation of the binding pocket of protein S100A10 was screened to identify compounds able to destabilize the complex between S100A10 and the phospholipid-binding protein Annexin A2 [64]. Steady-state FRET from Cy3-labeled Annexin A2 and Cy5-labeled S100A10 was employed in the screening. In this case, S100A10 protein labeling was not site directed, and its stoichiometry was only roughly defined. So, while fairly easily achieved, such an approach can only provide semiquantitative information on the efficacy of the tested small compounds in destabilizing the protein association.

FRET between two fluorescent probes, bound to the antibodies for two different epitope tags linked to the two monomers of HIV-1 integrase (IN), earlier used to characterize the monomer/dimer equilibrium of this protein [65], has later been combined with an equilibrium analysis of a binding model for IN–IN interaction, including the monomeric and several oligomeric species in the presence of an IN-dimer ligand. Dithiothreitol and β-mercaptoethanol weakened the IN monomer–monomer interaction. On the other hand, two peptides derived from LEDGF, a cellular cofactor that interacts with the IN-dimer interface, and a small molecule, all of which compete with LEDGF for binding to IN, were found to increase the stability of the IN dimers [66].

A nice example of the use of FRET, in combination with fluorescence microscopy, to monitor the effects of small compounds on PPIs in living cells is provided by the investigation of the effects of compounds, previously found able to disrupt BH3 interactions in vitro, on the heterodimerization of Bcl-XL with the pro-apoptotic proteins, BAX and BAD [46]. Intact cells were co-transfected with the expression vectors BAX fused to yellow fluorescent protein (YFP) and Bcl-XL fused to cyan fluorescent protein (CFP). Co-transfection resulted in an increase in the YFP-to-CFP emission ratio, relative to separately transfected cells, due to CFP-to-YFP energy transfer. The addition of the above compounds caused decreases in this ‘FRET ratio’ consistent with the activities of the compounds in vitro. The same approach, only involving different FRET donor and acceptor,

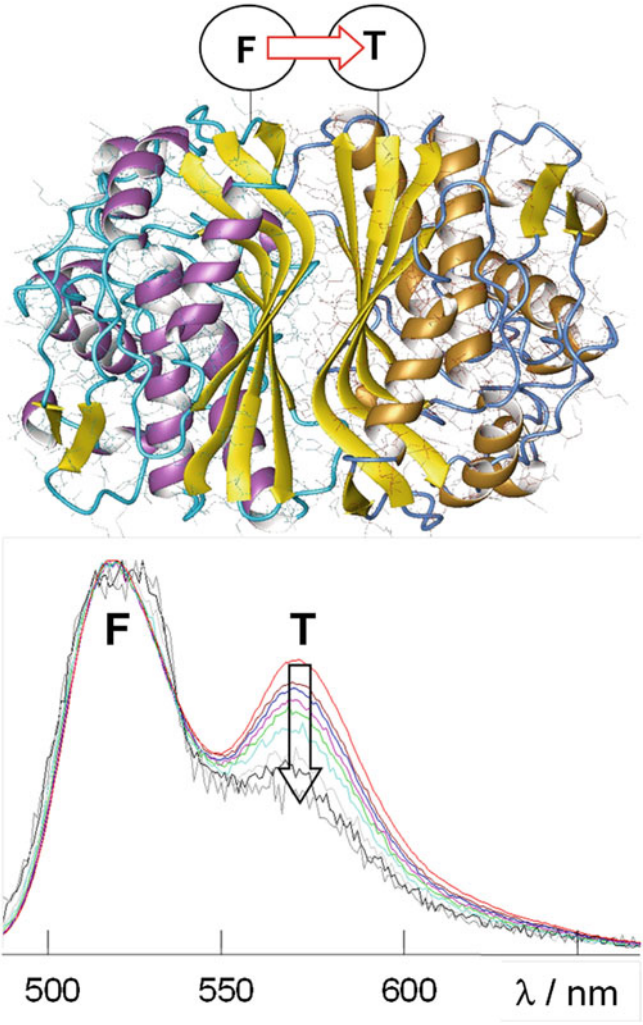


Fig. 6.3 FRET from fluorescein (*F*) to tetramethylrhodamine (*T*) bound to Cys 43 and 43' of the human thymidylate synthase dimer. FRET efficiency is correlated with the relative T/F emission intensity and decreases with decreasing total protein concentration, from ca 300 to ca 5 nM [62]

was employed to test the effects of the same compounds on heterodimerization of BAD with either Bcl-2 or Bcl-XL in intact cells.

Changes in the efficiency of homo-FRET, that is, excitation-energy transfer between a donor and an acceptor of the same chemical nature, such as two tryptophan residues or two extrinsic, identical probes in a protein oligomer, including homo-oligomers, may reveal changes in the protein oligomerization state [67]. Here, it is the depolarization associated with excitation-energy transfer between like fluorophores that is usually measured [28]. As a clever example of the use of

homo-FRET to investigate the effect of small molecules on protein oligomerization, we mention the case of the serotonin$_{1A}$ receptor whose oligomers are potentially implicated in the functional roles of the protein. Homo-FRET and fluorescence lifetime measurements have been used to monitor such an oligomerization in cells expressing the serotonin$_{1A}$ receptor tagged to enhanced yellow fluorescent protein [68]. The emission anisotropies were found to be lower than the value predicted for the monomeric protein, and the depolarization was attributed to homo-FRET within protein oligomers. To support this assignment, the fluorophores were progressively photobleached. Because the efficiency of homo-FRET correlates with the spectral absorption/emission overlap, the bleaching led to recovery of anisotropy. This was analyzed versus predicted recoveries for collections of dimers, trimers, and higher aggregates. This analysis, combined with the extrapolated anisotropies at full fluorophore photobleaching, enabled the authors to discriminate between oligomers of different sizes. In particular, they tested the effects of some known agonist and antagonists of the serotonin receptor and found that while treatment with an antagonist (p-MPPI) lowered the fraction of higher-order oligomers, the agonist (serotonin itself) induced the formation of higher-order oligomers. Homo-FRET combined with microscopy, and, in some instances, with the increased selectivity afforded by two-photon excitation, has recently provided subcellular resolution imaging of protein oligomers [69, 70].

Conventional, steady-state FRET measurements may suffer from limited accuracy because of interfering emissions, particularly from complex biological samples. An impressive increase in sensitivity has been obtained by employing long-lived fluorophores, mostly lanthanide ions, to enable time-gated measurements of the donor/acceptor signals delayed, with respect to excitation, from several microseconds to a few milliseconds, that is, a time when all background, usually nanosecond, emissions have decayed [71, 72]. Biosensors based on nanocrystals doped with lanthanides have been proposed for this application [72]. A similar increase in analytical robustness is obtained using a bioluminescent excitation-energy donor, typically, a luciferase (BRET, [13]). Here, intensities are extremely low, but interferences are essentially absent, as no excitation light is required. The methods are well suited for medium-to-high throughput screenings. An example of the first approach, often, and rather confusingly, called time-resolved FRET (TR-FRET), consists in the screening of 1,280 compounds to identify inhibitors of the dimerization of a 106-residue domain of the capsid protein of hepatitis C virus [73]. The Core-106 fragments were tagged with an N-terminal glutathione-S-transferase (GST) or Flag peptide. Europium cryptate, a long-lived donor, and allophycocyanin were used to label anti-GST and anti-Flag antibodies. Association between GST-core106 and Flag-core106 was assessed by measuring FRET between the two fluorophores following antigen/antibody recognition. Another example is provided by the discovery of potent, nonpeptide inhibitors of the interaction between leukocyte function-associated antigen-1 (LFA-1), a member of the β_2-integrin family of adhesion molecules, and intracellular adhesion molecule ICAM-1 [74, 75]. In this case the strategy consisted in immobilizing one of the two partners, tagged with a fluorescent probe, and adding the other partner, tagged with a long-lived europium

luminophore using the biotin/streptavidin recognition. The decrease in FRET observed in the presence of screened compounds was a measure of the ability of the latter to disrupt the LFA-1/ICAM-1 interaction. Several more HTS TR-FRET assays differing only in some details, say, the nature of the donating lanthanide and the accepting fluorophores or the strategy to tag the interacting proteins with the two fluorophores, have been applied to search for small PPI inhibitors. An Eu^{3+} cryptate-conjugated anti-FLAG antibody and an anti-6His antibody conjugated to a fluorescent excitation-energy acceptor have been employed in an assay designed to screen approximately 15,000 compounds to find inhibitors of the complexation of FLAG-fused IKKβ with NEMO-6His, a process involved in inflammatory and autoimmune disorders [76]. A Tb^{3+} chelate and Alexa Fluor 488 have been chemically conjugated, respectively, to the G protein, $G\alpha_{o,}$ and its regulator protein, RGS4, and used in a TR-FRET screening of approximately 40,000 compounds to find two inhibitors of this PPI [77]. An europium-labeled anti-His antibody and a streptavidin-conjugated APC fluorophores were employed to label the 6His–apoRBP4 and the biotinylated human TTR proteins to test small compounds that were found to either increase or decrease the affinity of the retinol-binding protein, RBP4, for transthyretin, TTR [78].

6.2.4 Multiple and Other Fluorescence Observables

The above-mentioned fluorescence observables may be usefully combined within the same investigation. An example is provided by the screening of 60 compounds that had been previously selected by computational methods as possible inhibitors of the down-regulation of the p53 tumor suppressor protein caused by interaction with the calcium-binding protein, S100B [79]. Complexation of the latter with the small compounds, leading to p53 activity increase, was monitored by four different titration experiments: direct measurement of emission changes from the fluorescent compounds due to subsequent additions of the S100B protein; quenching of tyrosine emission from the protein upon titration with the compounds; measurement of tryptophan emission restoration in competition titrations of wild-type S100B into solutions of the complexes of an S100B tryptophan mutant with the small compounds; and measurement of fluorescence from a peptide derived from p53 (F385W) that binds holo-S100B in competition titrations of the small molecules to the S100B-p53 complex.

Another example of combination of fluorescence observables is provided by the screening of small-molecule inhibitors that interfere with the cytohesin-catalyzed GDP/GTP exchange on a truncated version of ARF1, an adenosine diphosphate ribosylation factor (NΔ17ARF1) and/or with the interaction between NΔ17ARF1-GTP and its effector protein GGA3 [80]. The two proteins were fused to the fluorescent proteins CyPet and YPet, respectively. To identify the two kinds of inhibitors, the nucleotide exchange on NΔ17ARF1 was monitored in real time by measuring the associated enhancement of its intrinsic tryptophan fluorescence,

while association of the two proteins was simultaneously monitored by measuring the CyPet-to-YPet FRET: the two phenomena increased with similar rates, suggesting GDP/GTP exchange to be rate limiting. As often found, applications of the first assay were limited to the tested compounds that did not act as inner filters for tryptophan excitation, that is, that absorbed negligibly at 280 nm.

Among the fluorescence observables, some have not been employed to investigate the action of small molecules able to interfere with PPIs. A remarkable example is provided by fluorescence cross-correlation spectroscopy (FCCS, [81]). FCCS is a powerful tool to monitor protein–protein and protein–DNA interactions, both in solution and in cells. Measurement of cross correlation between the time-fluctuations of the emissions from two different, independently excited probes, each attached to a partner of the interacting pair (or larger assembly), provides information on the complex dynamics and thermodynamics. The FCCS approach suffers from difficulties related with probe binding to the interacting partners, especially for monitoring in cells. The problem, however, has now been solved in many cases by employing different strategies, including autofluorescent labeling, that is, expression of the protein of interest fused with different fluorescent proteins, specific chemical labeling, use of fluorescent antibodies [81], or by employing two-photon excited intrinsic protein fluorescence [82]. The method, now implemented on commercial fluorescence microscopes, is therefore recommended for monitoring perturbation of PPIs by small molecules.

While TR emission from probes has found applications (some are quoted in the previous paragraphs), measurement of the time-course of intrinsic protein emission to monitor changes in the aggregation state is apparently an unexplored opportunity. There are a number of practical reasons that make this kind of experiments poorly apt for medium/HTS: Instrumentation is often expensive, measurements are usually time-consuming, and analysis of the results may be rather complex [83]. While these observations are intimately related with the interactions experienced by tyrosine and tryptophan residues in the protein, and probably reflect even subtle structural changes with an unprecedented sensitivity, a structural/dynamic interpretation of the changes observed in the time-course of a protein emission remains a difficult task (Ref. [29], Chap. 17). However, because of the wealth of information buried within, efforts have been made, and are currently underway, to set up tools and knowledge able to extract this information [33, 83, 84]. Therefore, it is easy to predict TR intrinsic protein fluorescence to become a major source of structural/mechanistic information on PPIs and their modulation.

6.3 Dissociative Inhibition Kinetics

While not directly monitoring PPI perturbations by small molecules, whenever a catalytic efficiency depends on some protein multimeric assembly, kinetic analysis can provide direct evidence of the mechanistic consequences of such perturbations. The key observation in these studies is a modulation by small ligands of the

dependence of specific enzyme activity on protein concentration, that is associated with a mass-balance-law-governed distribution of the protein monomers and various multimers. This effect of added small molecules results from their interfering with the interactions between enzyme subunits usually leading to destabilization of the multimeric assembly.

A few different inhibition models of multimeric enzymes have been proposed that involve destabilization of protein–protein attractive interactions. The inhibitor may bind to some protein sequence that is only exposed during the folding process and thus prevents protein association during folding. In an example of this kind, peptides mimicking one or two β-strands from the human immunodeficiency virus 1 (HIVl) interface were shown to inhibit the dimeric enzymes, HIVl and HIV2 proteases [85]. A standard kinetic analysis indicated a noncompetitive inhibition mechanism, with, however, no hint at the dimeric nature of the enzyme, or at the possibly dissociative character of the inhibition.

For dimeric enzymes, the analysis of the so-called dissociative inhibition model has been provided in Ref. [86]. An inhibitor of a functionally obligate dimeric enzyme was assumed to bind the dimer (competitive inhibition), the monomer (dissociative inhibition), the dimer-substrate complex (uncompetitive inhibition), or both the dimer and the dimer-substrate complex (noncompetitive inhibition, Fig. 6.4). Resolution of the kinetic scheme in the rapid equilibrium regime led to the expectation that $E_0/\sqrt{k_{exp}}$ versus $\sqrt{k_{exp}}$ plots—E_0 being the total enzyme concentration that was varied in the experiments and k_{exp} the ratio of the initial reaction rate and the total substrate concentration, which was kept constant—were linear with constant slopes and increasing intercepts at increasing inhibitor concentrations ('Zhang–Poorman plots'). From the best-fit slopes and intercepts, the relevant parameters of the kinetic model were obtained, including the affinities of the inhibitor for the monomeric and the dimeric enzyme, K_I and K_C, respectively, and the monomer/dimer dissociation constant of the enzyme, K_D. The authors applied their analysis to demonstrate that a tetrapeptide corresponding to the COOH terminal segment of HIV-1 protease was, indeed, a dissociative inhibitor, that is, it bound to the inactive monomers (M) and prevented their association into the active dimer (M_2).

In the original model, the inhibition was studied under first-order conditions, that is, the total substrate concentration, [S], was assumed much smaller than K_M. An alternative solution of the kinetic scheme has been recently obtained without making this assumption, in order to extend the analysis to cases in which fulfillment of this condition would require very small [S] values and the need to measure prohibitively slow reaction kinetics [18].

In the recent literature, conformation of kinetic data to this model, as judged from the linearity of the Zhang–Poorman plots and a significant dependence of the intercepts on inhibitor concentrations, has been shown in a number of examples, many of which concern HIV1 protease dissociative inhibitors. To mention a few, dissociative inhibition was found with some nine-residue peptides obtained through an impressive genetic-selection approach [87], with a 27-residue peptide designed from domains at the N- and C-termini of the same enzyme [88], with

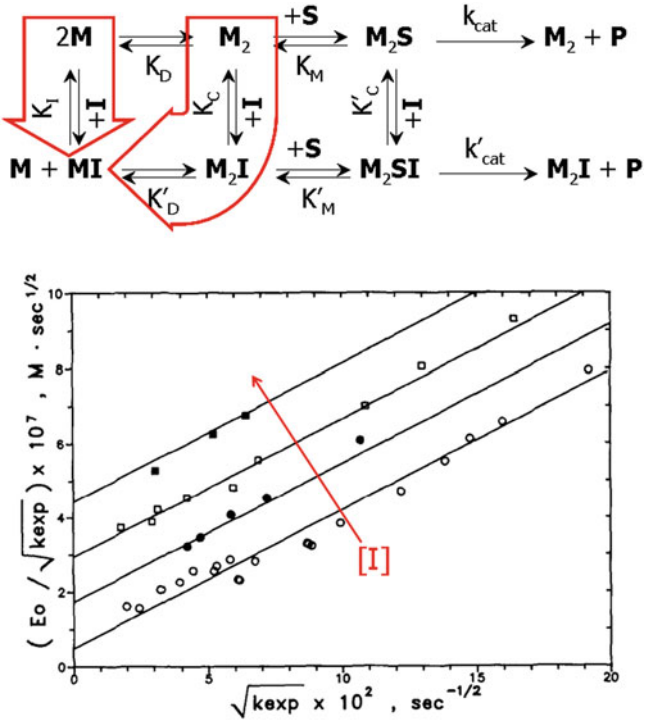


Fig. 6.4 *Top* kinetic scheme for inhibition of an obligatory dimeric enzyme (adapted from Ref. [86]). Dissociative inhibition is represented by the $M \rightarrow MI$ and/or the $M_2 \rightarrow M_2I \rightarrow MI$ paths (*red large arrows*). *Bottom* the dissociative inhibition fingerprint: $E_0 \sqrt{k_{exp}}$ versus $\sqrt{k_{exp}}$ plots are parallel lines with intercepts that increase with increasing inhibitor concentration, and the [I] = 0 line has a non-null intercept (taken from Ref. [86]). The meanings of the symbols are given in the text

some interface peptides cross-linked at their amino termini [89, 90] and at side chains [91], with tetracyclic triterpene schisanlactone, a natural product isolated from a fungus [92], and with some naphthalene- and quinoline-based nonpeptidic 'molecular tongs' [93–95]. The above kinetic analysis was corroborated by analytical ultracentrifugation results to characterize some alkyl tripeptides as dissociative inhibitors of the same enzyme, both wild type and mutated [96]. Dissociative inhibition seems to be a useful strategy also versus two other HIV-1 enzymes, reverse transcriptase and IN [97].

Among the fewer examples concerning other dimeric enzymes, we mention the inhibition of dimeric aminoimidazole carboxamide ribonucleotide transformylase (AICAR Tfase) by a small compound, Cappsin 1 [98], and that of 3C-like proteinase of severe acute respiratory syndrome coronavirus by some octapeptides derived from the protein N-terminal [99]. Apparently, no available examples involve more-than-dimeric enzymes, and the dissociative inhibition model has not been extended to higher oligomers than dimers.

Of some interest is the mechanistic issue concerning whether a supposed inhibitor of a dimeric enzyme preferentially binds the enzyme monomer, thus preventing its association with another monomer to obtain the active dimer, or binds the already formed dimer causing its disruption. The two paths may require quite different molecular properties for an efficient inhibitor, and, as a consequence, different molecular design strategies. The kinetic model in Fig. 6.4 encompasses both mechanistic routes, that we may simplify as, respectively, $M \rightarrow MI$ and $M_2 \rightarrow M_2I \rightarrow MI$, and characterizes them through the corresponding equilibrium constants, K_I and K_CK_D'. Should one path be much slower than the other and the corresponding equilibration not attained in the experimental runs, a more sophisticated kinetic analysis involving all the relevant rate constants and more detailed TR experimental information would likely be necessary to conclude which of the two kinetic paths proposed remains the only functionally relevant one.

Lastly, we remark that dissociative inhibition is not the only mechanism by which a small molecule can inhibit a multimeric obligate enzyme. Kinetic analysis, in combination with crystallographic and calorimetric evidence, showed that inhibition of *Trypanosoma cruzi* triosephosphate isomerase was caused by a small molecule whose binding triggered evolution of the dimeric protein toward an inactive conformation, rather than to dimer disruption. The nonlinear dependence of pseudo-first-order constants of inactivation on inhibitor concentration provided information on the complex inhibition mechanism [100]. As briefly reported in paragraph 2.3, use of FRET between probes enabled the authors to rule out a dissociative mechanism to interpret inhibition of dimeric human TS by some octapeptides designed from a sequence in the inter-monomer surface [63]. Experimental evidence, kinetic, crystallographic, and calorimetric, led to the conclusion that the peptides selectively bound to a dimeric, inactive conformation of the protein, thus stabilizing it. A specific kinetic scheme was solved under the usual fast equilibrium assumption and was shown to be consistent with the observed noncompetitive kinetic behavior with, however, modified interpretations of the plot slopes, intercepts, and crossing point.

Acknowledgements This work supported by the Italian Association for Cancer Research (AIRC), grant AIRC-DROC IG 10474.

References

1. Weiss A, Schlessinger J (1998) Switching signals on or off by receptor dimerization. Cell 94:277–280
2. Fry DC, Vassilev LT (2005) Targeting protein–protein interactions for cancer therapy. J Mol Med 83:955–963
3. Stites WE (1997) Protein-protein interactions: interface structure, binding thermodynamics, and mutational analysis. Chem Rev 97:1233–1250
4. Block P, Weskamp N, Wolf A, Klebe G (2007) Strategies to search and design stabilizers of protein–protein interactions: a feasibility study. Proteins: Struct, Funct, Bioinf 681:70–186

5. Silvian LF, Friedman JE, Strauch K, Cachero TG, Day ES, Qian F, Cunningham B, Fung A, Sun L, Su L, Zheng Z, Kumaravel G, Whitty A (2011) Small molecule inhibition of the TNF family cytokine CD40 ligand through a subunit fracture mechanism. ACS Chem Biol 6:636–647
6. Wilson CGM, Arkin MR (2011) Small-molecule inhibitors of IL-2/IL-2R: lessons learned and applied. In: Vassiliev L, Fry D (eds) Small-molecule inhibitors of protein–protein interactions. Curr Top Microbiol, vol. 348. Springer, Berlin
7. Thangudu RR, Bryant SH, Panchenko AR, Madej T (2012) Modulating protein–protein interactions with small molecules: the importance of binding hotspots. J Mol Biol 415:443–453
8. Berg T (2003) Modulation of protein–protein interactions with small organic molecules. Angew Chem Int Ed 42:2462–2481
9. Braisted AC, Oslob JD, Delano WL, Hyde J, McDowell RS, Waal N, Yu C, Arki MR, Raimundo BC (2003) Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. J Am Chem Soc 125:3714–3715
10. Arkin MR, Wells JA (2004) Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. Nat Rev Drug Discov 3:301–317
11. Phizicky EM, Fields S (1995) Protein-protein interactions: methods for detection and analysis. Microbiol Rev 59(1):94–123
12. Lakey JH, Raggett EM (1998) Measuring protein–protein interactions. Curr Opin Struc Biol 8:119–123
13. Piehler J (2005) New methodologies for measuring protein interactions in vivo and in vitro. Curr Opin Struc Biol 15:4–14
14. Shoemaker BA, Panchenko AR (2007) Deciphering protein–protein interactions. Part I. Experimental techniques and databases. PLoS Comput Biol 3:337–344
15. Heeres JT, Hergenrother PJ (2011) High-throughput screening for modulators of protein–protein interactions: use of photonic crystal biosensors and complementary technologies. Chem Soc Rev 40:4398–4410
16. Vaynberg J, Qin J (2006) Weak protein–protein interactions as probed by NMR spectroscopy. Trends Biotechnol 24:22–27
17. Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F (2007) Aggregation analysis of therapeutic proteins. Part 2. Analytical ultracentrifugation and dynamic light scattering. Bioprocess Int 5:36–47. http://www.ap-lab.com/Part2.pdf
18. Cardinale D, Salo-Ahen OMH, Ferrari S, Ponterini G, Cruciani G, Carosati E, Tochowicz AM, Mangani S, Wade RC, Costi MP (2010) Homodimeric enzymes as drug targets. Curr Med Chem 17:826–846
19. Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F (2006) Aggregation analysis of therapeutic proteins. Part 1. General aspects and techniques for assessment. Bioprocess Int 4:42–43. http://www.ap-lab.com/Bioprocess_Intl_Aggregation_Part1.pdf
20. Patapoff TW, Mrsny RJ, Lee WA (1993) The application of size exclusion chromatography and computer simulation to study the thermodynamic and kinetic parameters for short-lived dissociable protein aggregates. Anal Biochem 212:71–78
21. Hall D, Huang L (2012) On the use of size exclusion chromatography for the resolution of mixed amyloid aggregate distributions: I. equilibrium partition models. Anal Biochem 426:69–85
22. Karlsson R (2004) SPR for molecular interaction analysis: a review of emerging application areas. J Mol Recognit 17:151–161
23. Raimundo BC, Oslob JD, Braisted AC, Hyde J, McDowell RS, Randal M, Waal ND, Wilkinson J, Yu CH, Arkin MR (2004) Integrating fragment assembly and biophysical methods in the chemical advancement of small-molecule antagonists of IL-2: an approach for inhibiting protein–protein interactions. J Med Chem 47:3111–3130
24. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303:844–848

25. Van der Merwe PA (2000) Surface plasmon resonance. In: Harding S, Chowdhry PZ (eds) Protein–ligand interactions: a practical approach. Oxford University Press, New York
26. Zuiderweg ERP (2002) Mapping protein–protein interactions in solution by NMR spectroscopy. Biochemistry 41:1–7
27. Morris MA (2010) Fluorescent biosensors of intracellular targets from genetically encoded reporters to modular polypeptide probes. Cell Biochem Biophys 56:19–37
28. Valeur B (2002) Molecular fluorescence. Principles and applications. Wiley-VCH, Weinheim
29. Lakowicz JR (2006) Principles of fluorescence spectroscopy. Springer, New York
30. Otto-Bruc A, Antonny B, Vuong TM, Chardin P, Chabré M (1993) Interaction between the retinal cyclic GMP phosphodiesterase inhibitor and transducin. Kinetics and affinity studies. Biochemistry 32:8636–8645
31. Marion JD, Van DN, Bell JE, Bell JK (2010) Measuring the effect of ligand binding on the interface stability of multimeric proteins using dynamic light scattering. Anal Biochem 407:278–280
32. Zutshi R, Franciskovich J, Shultz M, Schweitzer B, Bishop P, Wilson M, Chmielewski J (1997) Targeting the dimerization interface of HIV-1 protease: inhibition with cross-linked interfacial peptides. J Am Chem Soc 119:4841–4845
33. Engelborghs Y (2001) The analysis of time resolved protein fluorescence in multi-tryptophan proteins. Spectrochim Acta A 57:2255–2270
34. Maithal K, Ravindra G, Nagaraj G, Kumar Singh S, Balaram H, Balaram P (2002) Subunit interface mutation disrupting an aromatic cluster in *Plasmodium falciparum* triosephosphate isomerase: effect on dimer stability. Prot Eng 15:575–584
35. Garai K, Frieden C (2010) The association-dissociation behavior of the ApoE proteins: kinetic and equilibrium studies. Biochemistry 49:9533–9541
36. Kipp RA, Case MA, Wist AD, Cresson CM, Carrell M, Griner E, Wiita A, Albiniak PA, Chai J, Shi Y, Semmelhack MF, McLendon GL (2002) Molecular targeting of inhibitor of apoptosis proteins based on small molecule mimics of natural binding partners. Biochemistry 41:7344–7349
37. Kersten S, Pan L, Chambon P, Gronemeyer H, Nay N (1995) Role of ligand in retinoid signaling. 9-cis-Retinoic acid modulates the oligomeric state of the retinoid X receptor. Biochemistry 34:13717–13721
38. Khanna M, Chelladurai B, Gavini A, Li L, Shao M, Courtney D, Turchi JJ, Matei D, Meroueh S (2011) Targeting ovarian tumor cell adhesion mediated by tissue transglutaminase. Mol Cancer Ther 10:626–636
39. Liu Z, Sun C, Olejniczak ET, Meadows RP, Betz SF, Oost T, Herrmann J, Wu JC, Fesik SW (2000) Structural basis for binding of Smac/DIABLO to the XIAP BIR3 domain. Nature 408:1004–1008
40. Sun H, Nikolovska-Coleska Z, Yang C-Y, Xu L, Liu M, Tomita Y, Pan H, Yoshioka Y, Krajewski K, Roller PP, Wang S (2004) Structure-based design of potent conformationally constrained Smac mimetics. J Am Chem Soc 126:16686–16687
41. Oost TK, Sun C, Armstrong RC, Al-Assaad A-S, Betz SF, Deckwerth TL, Ding H, Elmore SW, Meadows RP, Olejniczak ET, Oleksijew A, Oltersdorf T, Rosenberg SH, Shoemaker AR, Tomaselli KJ, Zou H, Fesik SW (2004) Discovery of potent antagonists of the antiapoptotic protein XIAP for the treatment of cancer. J Med Chem 47:4417–4426
42. Buckley DL, Van Molle I, Gareiss PC, Tae HS, Michel J, Noblin DJ, Jorgensen WL, Ciulli A, Crews CM (2012) Targeting the von Hippel–Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1α interaction. J Am Chem Soc 134:4465–4468
43. Leone M, Zhai D, Saret S, Kitada S, Reed JC, Pellecchia M (2003) Cancer prevention by tea polyphenols is linked to their direct inhibition of antiapoptotic Bcl-2-family proteins. Cancer Res 63:8118–8121
44. Wang J-L, Liu D, Zhang Z-J, Shan S, Han X, Srinivasula SM, Croce CM, Alnemri ES, Huang Z (2000) Structure-based discovery of an organic compound that binds Bcl-2 protein and induces apoptosis of tumor cells. P Natl Acad Sci USA 97:7124–7129

45. Enyedy IJ, Ling Y, Nacro K, Tomita Y, Wu X, Cao Y, Guo R, Li B, Zhu X, Huang Y, Long Y-Q, Roller PP, Yang D, Wang S (2001) Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. J Med Chem 44:4313–4324
46. Degterev A, Lugovskoy A, Cardone M, Mulley B, Wagner G, Mitchison T, Yuan J (2001) Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-XL. Nat Cell Biol 3:173–182
47. Chan S-L, Lee MC, Tan KO, Yang L-K, Lee ASY, Flotow H, Fu NY, Butler MS, Soejarto DD, Buss AD, Yu VC (2003) Identification of chelerythrine as an inhibitor of Bcl-XL function. J Biol Chem 278:20453–20456
48. Yin H, Hamilton AD (2004) Terephthalamide derivatives as mimetics of the helical region of BAK peptide target Bcl-XL protein. Bioorg Med Chem Lett 14:1375–1379
49. Rush TS, Grant JA, Mosyak L, Nicholls A (2005) A shape-based 3-D scaffold hopping method and its application to a bacterial protein–protein interaction. J Med Chem 48:1489–1495
50. Zhao J, Du Y, Horton JR, Upadhyay AK, Lou B, Bai Y, Zhang X, Du L, Li M, Wang B, Zhang L, Barbieri JT, Khuri FR, Cheng X, Fu H (2011) Discovery and structural characterization of a small molecule 14–3-3 protein–protein interaction inhibitor. P Natl Acad Sci USA 108:16212–16216
51. Caselli M, Latterini L, Ponterini G (2004) Consequences of H-dimerization on the photophysics and photochemistry of oxacarbocyanines. Phys Chem Chem Phys 6:3857–3863
52. He MM, Stroustrup Smith A, Oslob JD, Flanagan WM, Braisted AC, Whitty A, Cancilla MT, Wang J, Lugovskoy AA, Yoburn JC, Fung AD, Farrington G, Eldredge JK, Day ES, Cruz LA, Cachero TG, Miller SK, Friedman JE, Choong IC, Cunningham BC (2005) Small-molecule inhibition of TNF-α. Science 310:1022–1025
53. Prasanna V, Bhattacharjya S, Balaram P (1998) Synthetic interface peptides as inactivators of multimeric enzymes: inhibitory and conformational properties of three fragments from *Lactobacillus casei* thymidylate synthase. Biochemistry 37:6883–6893
54. Grasberger BL, Lu T, Schubert C, Parks DJ, Carver TE, Koblish HK, Cummings MD, LaFrance LV, Milkiewicz KL, Calvo RR, Maguire D, Lattanze J, Franks CF, Zhao S, Ramachandren K, Bylebyl GR, Zhang M, Manthey CL, Petrella EC, Pantoliano MW, Deckman IC, Spurlino JC, Maroney AC, Tomczuk BE, Molloy CJ, Bone RF (2005) Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. J Med Chem 48:909–912
55. Tzung S-P, Kim KM, Basañez G, Giedt CD, Simon J, Zimmerberg J, Zhang KYJ, Hockenbery DM (2001) Antimycin A mimics a cell-death-inducing Bcl-2 homology domain 3. Nat Cell Biol 3:183–191
56. Kim KM, Giedt CD, Basañez G, O'Neill JW, Hill JJ, Han Y-H, Tzung S-P, Zimmerberg J, Hockenbery DM, Zhang KYJ (2001) Biophysical characterization of recombinant human Bcl-2 and its interactions with an inhibitory ligand, antimycin A. Biochemistry 40:4911–4922
57. Koivunen Nissinen L, Käpylä J, Jokinen J, Pihlavisto M, Marjamäki A, Heino J, Huuskonen J, Pentikäinen OT (2011) Fluorescent small molecule probe to modulate and explore $\alpha 2\beta 1$ integrin function. J Am Chem Soc 133:14558–14561
58. Cosa G, Focsaneanu K-S, McLean JRN, McNamee JP, Scaiano JC (2001) Photophysical properties of fluorescent DNA-dyes bound to single- and double-stranded DNA in aqueous buffered solution. Photochem Photobiol 73:585–599
59. Visser AJWG, Laptenok SP, Visser NV, van Hoek A, Birch DJS, Brochon J-C, Borst JW (2010) Time-resolved FRET fluorescence spectroscopy of visible fluorescent protein pairs. Eur Biophys J 39(2):241–253
60. Truong K, Ikura M (2001) The use of FRET imaging microscopy to detect protein–protein interactions and protein conformational changes in vivo. Curr Opin Struc Biol 11:573–578
61. Li HY, Ng EKO, Lee SMY, Kotaka M, Tsui SKW, Lee CY, Fung KP, Waye MMY (2001) Protein-protein interaction of FHL3 with FHL2 and visualization of their interaction by

green fluorescent proteins (GFP) two-fusion fluorescence resonance energy transfer (FRET). J Cell Biochem 80:293–303
62. Genovese F, Ferrari S, Guaitoli G, Caselli M, Costi MP, Ponterini G (2010) Dimer–monomer equilibrium of human thymidylate synthase monitored by fluorescence resonance energy transfer. Protein Sci 19:1023–1030
63. Cardinale D, Guaitoli G, Tondi D, Luciani R, Henrich S, Salo-Ahen OMH, Ferrari S, Marverti G, Guerrieri D, Ligabue A, Frassineti C, Pozzi C, Mangani S, Fessas D, Guerrini R, Ponterini G, Wade RC, Costi MP (2011) Protein-protein interface-binding peptides with a novel inhibitory mechanism against the cancer target, human thymidylate synthase. P Natl Acad Sci USA 108:E542–E549
64. Li C, Reddy TRK, Fischer PM, Dekker LV (2010) A Cy5-labeled S100A10 tracer used to identify inhibitors of the protein interaction with annexin A2. Assay Drug Dev Techn 8:85–95
65. Tsiang M, Jones GS, Hung M, Mukund S, Han B, Liu X, Babaoglu K, Lansdon E, Chen X, Todd J, Cai T, Pagratis N, Sakowicz R, Geleziunas R (2009) Affinities between the binding partners of the HIV-1 integrase dimer/lens epithelium derived growth factor (IN dimer/LEDGF) complex. J Biol Chem 284:33580–33599
66. Tsiang M, Jones GS, Hung M, Samuel D, Novikov N, Mukund S, Brendza KM, Niedziela-Majka A, Jin D, Liu X, Mitchell M, Sakowicz R, Geleziunas R (2011) Dithiothreitol causes HIV-1 integrase dimer dissociation while agents interacting with the integrase dimer interface promote dimer formation. Biochemistry 50:1567–1581
67. Levitt JA, Matthews DR, Ameer-Beg SM, Suhling K (2009) Fluorescence lifetime and polarization-resolved imaging in cell biology. Curr Opin Biotech 20:28–36
68. Ganguly S, Clayton AHA, Chattopadhyay A (2011) Organization of higher-order oligomers of the serotonin$_{1A}$ receptor explored utilizing homo-FRET in live cells. Biophys J 100:361–368
69. Bader AN, Hofman EG, Voortman J, Gerritsen HC, van Bergen En Henegouwen PMP (2009) Homo-FRET imaging enables quantification of protein cluster sizes with subcellular resolution. Biophys J 97(9):2613–2622
70. Hofman EG, Bader AN, Voortman J, van den Heuvel DJ, Sigismund S, Verkleij AJ, Gerritsen HC, van Bergen en Henegouwen PMP (2010) Ligand-induced EGF receptor oligomerization is kinase-dependent and enhances internalization. J Biol Chem 285:39481–39489
71. Hemmila I, Laitala V (2005) Progress in lanthanides as luminescent probes. J Fluoresc 15:529–542
72. Tu D, Liu L, Ju Q, Liu Y, Zhu H, Li R, Chen X (2011) Time-resolved FRET biosensor based on amine-functionalized lanthanide-doped $NaYF_4$ nanocrystals. Angew Chem Int Ed 50:6306–6310
73. Kota S, Scampavia L, Spicer T, Beeler AB, Takahashi V, Snyder JK, Porco JA Jr, Hodder P, Strosberg AD (2010) A time-resolved fluorescence–resonance energy transfer assay for identifying inhibitors of hepatitis C virus core dimerization. Assay Drug Dev Techn 8:96–105
74. Liu G, Link JT, Pei Z, Reilly EB, Leitza S, Nguyen B, Marsh KC, Okasiski GF, von Geldern TW, Ormes M, Fowler K, Gallatin M (2000) Discovery of novel p-arylthio cinnamides as antagonists of leukocyte function-associated antigen-1/intracellular adhesion molecule-1 interaction. 1. Identification of an addition binding pocket based on an anilino diaryl sulfide lead. J Med Chem 43:4025–4040
75. Pei Z, Xin Z, Liu G, Li Y, Reilly EB, Lubbers NL, Huth JR, Link JT, von Geldern TW, Cox BF, Leitza S, Gao Y, Marsh KC, DeVries P, Okasinski GF (2001) Discovery of potent antagonists of leukocyte function-associated antigen-1/intercellular adhesion molecule-1 interaction. 3.Amide (C-ring) structure-activity relationship and improvement of overall properties of arylthio cinnamides. J Med Chem 44:2913–2920
76. Gotoh Y, Nagata H, Kase H, Shimonishi M, Ido M (2010) A homogeneous time-resolved fluorescence-based high-throughput screening system for discovery of inhibitors of IKKβ–NEMO interaction. Anal Biochem 405:19–27

77. Blazer LL, Roman DL, Chung A, Larsen MJ, Greedy BM, Husbands SM, Neubig RR (2010) Reversible, allosteric small-molecule inhibitors of regulator of G protein signaling proteins. Mol Pharmacol 78:524–533
78. Coward P, Conn M, Tang J, Xiong F, Menjares A, Reagan JD (2009) Application of an allosteric model to describe the interactions among retinol binding protein 4, transthyretin, and small molecule retinol binding protein 4 ligands. Anal Biochem 384:312–320
79. Markowitz J, Chen I, Gitti R, Baldisseri DM, Pan Y, Udan R, Carrier F, MacKerell AD, Weber DJ (2004) Identification and characterization of small molecule inhibitors of the calcium-dependent S100B–p53 tumor suppressor interaction. J Med Chem 47:5085–5093
80. Bill A, Blockus H, Stumpfe D, Bajorath J, Schmitz A, Famulok M (2011) A homogeneous fluorescence resonance energy transfer system for monitoring the activation of a protein switch in real time. J Am Chem Soc 133:8372–8379
81. Langowski J (2008) Protein–protein interactions determined by fluorescence correlation spectroscopy. Method Cell Biol 85:471–484
82. Sahoo B, Balaji J, Nag S, Kaushalya SK, Maiti S (2008) Protein aggregation probed by two-photon fluorescence correlation spectroscopy of native tryptophan. J Chem Phys 129:075103-1–075103-5
83. Lee JC, Langen R, Hummel PA, Gray HB, Winkler JR (2004) α-Synuclein structures from fluorescence energy-transfer kinetics: implications for the role of the protein in Parkinson's disease. P Natl Acad Sci USA 101:16466–16471
84. Engelborghs Y (2003) Correlating protein structure and protein fluorescence. J Fluoresc 13:9–16
85. Babé LM, Rose J, Craik CS (1992) Synthetic "interface" peptides alter dimeric assembly of the HIV 1 and 2 proteases. Protein Sci 1:1244–1253
86. Zhang Z-Y, Poorman RA, Maggiora LL, Heinrikson RL, Kezdy FJ (1991) Dissociative inhibition of dimeric enzymes. J Biol Chem 266:15591–15594
87. Park S-H, Raines RT (2000) Genetic selection for dissociative inhibitors of designated protein–protein interactions. Nat Biotech 18:847–851
88. Davis DA, Brown CA, Singer KE, Wang V, Kaufman J, Stahl SJ, Wingfield P, Maedaa K, Harada S, Yoshimura K, Kosalaraksa P, Mitsuya H, Yarchoan R (2006) Inhibition of HIV-1 replication by a peptide dimerization inhibitor of HIV-1 protease. Antivir Res 72:89–99
89. Hwang YS, Chmielewski J (2005) Development of low molecular weight HIV-1 protease dimerization inhibitors. J Med Chem 48:2239–2242
90. Lee S-G, Chmielewski J (2006) Rapid synthesis and in situ screening of potent HIV-1 protease dimerization inhibitors. Chem Biol 13:421–426
91. Bowman MJ, Chmielewski J (2009) Sidechain-linked inhibitors of HIV-1 protease dimerization. Bioorg Med Chem 17:967–976
92. El Dine RS, El Halawany AM, Ma C-M, Hattori M (2009) Inhibition of the dimerization and active site of HIV-1 protease by secondary metabolites from the vietnamese mushroom Ganoderma colossum. J Nat Prod 72:2019–2023
93. Merabet N, Dumond J, Collinet B, Van Baelinghem L, Boggetto N, Ongeri S, Ressad F, Reboud-Ravaux M, Sicsic S (2004) New constrained "molecular tongs" designed to dissociate HIV-1 protease dimer. J Med Chem 47:6392–6400
94. Bannwarth L, Kessler A, Pèthe S, Collinet B, Merabet N, Boggetto N, Sicsic S, Reboud-Ravaux M, Ongeri S (2006) Molecular tongs containing amino acid mimetic fragments: new inhibitors of wild-type and mutated HIV-1 protease dimerization. J Med Chem 49:4657–4664
95. Vidu A, Dufau L, Bannwarth L, Soulier J-L, Sicsic S, Piarulli U, Reboud-Ravaux M, Ongeri S (2010) Toward the first nonpeptidic molecular tong inhibitor of wild-type and mutated HIV-1 protease dimerization. Chem Med Chem 5:1899–1906
96. Bannwarth L, Rose T, Dufau L, Vanderesse R, Dumond J, Jamart-Grégoire B, Pannecouque C, De Clercq E, Reboud-Ravaux M (2009) Dimer disruption and monomer sequestration by alkyl tripeptides are successful strategies for inhibiting wild-type and multidrug-resistant mutated HIV-1 proteases. Biochemistry 48:379–387

97. Camarasa M-J, Velàzquez S, San-Fèlix A, Pérez-Pérez MJ, Gago F (2006) Dimerization inhibitors of HIV-1 reverse transcriptase, protease and integrase: a single mode of inhibition for the three HIV enzymes? Antivir Res 71:260–267
98. Capps KJ, Humiston J, Dominique R, Hwang I, Boger DL (2005) Discovery of AICAR Tfase inhibitors that disrupt requisite enzyme dimerization. Bioorg Med Chem Lett 15:2840–2844
99. Wei P, Fan K, Chen H, Ma L, Huang C, Tan L, Xi D, Li C, Liu Y, Cao A, Lai L (2006) The N-terminal octapeptide acts as a dimerization inhibitor of SARS coronavirus 3C-like proteinase. Biochem Bioph Res Co 339:865–872
100. Téllez-Valencia A, Olivares-Illana V, Hernàndez-Santoyo A, Pèrez-Montfort R, Costas M, Rodrìguez-Romero A, Lòpez-Calahorra F, Tuena de Gòmez-Puyou M, Gòmez-Puyou A (2004) Inactivation of triosephosphate isomerase from *Trypanosoma cruzi* by an agent that perturbs its dimer interface. J Mol Biol 341:1355–1365

Index

S. Mangani (ed.), *Disruption of Protein–Protein Interfaces*,
DOI: 10.1007/978-3-642-37999-4, © Springer-Verlag Berlin Heidelberg 2013