Protein Engineering, Design & Selection vol. 24 no. 9 pp. 635–648, 2011
Published online June 15, 2011 doi:10.1093/protein/gzr025

Transient protein–protein interactions

Saliha Ece Acuner Ozbabacan, Hatice Billur Engin, Attila Gursoy[1] and Ozlem Keskin[1]

Center for Computational Biology and Bioinformatics, College of Engineering, Koc University, Rumelifeneri Yolu, 34450 Sariyer Istanbul, Turkey

[1]To whom correspondence should be addressed.
E-mail: okeskin@ku.edu.tr (O.K.) and agursoy@ku.edu.tr (A.G.)

Received February 24, 2010; revised May 11, 2011; accepted May 17, 2011

Edited By Frederic Rousseau

Transient complexes are crucial for diverse biological processes such as biochemical pathways and signaling cascades in the cell. Here, we give an overview of the transient interactions; the importance of transient interactions as drug targets; and the structural characterization of transient protein–protein complexes based on the geometrical and physicochemical features of the transient complexes' interfaces. To better understand and eventually design transient protein–protein interactions (TPPIs), a molecular perspective of the protein–protein interfaces is necessary. Obtaining high-quality structures of protein–protein interactions could be one way of achieving this goal. After introducing the association kinetics of TPPIs, we elaborate on the experimental techniques detecting TPPIs in combination with the computational methods which classify transient and/or non-obligate complexes. In this review, currently available databases and servers that can be used to identify and predict TPPIs are also compiled.

Keywords: protein interface/protein interaction types/ transient protein − protein interactions/transient protein–drug interactions

Introduction

Transient complexes are essential, especially in the regulation of biochemical pathways and signaling cascades in the cells. A wide range of biological processes, such as hormone–receptor binding, signal transduction, allostery of enzymes, inhibition of proteases and correction of misfolded proteins by chaperones contain transient interactions between proteins (Valdar and Thornton, 2001; Schreiber *et al.*, 2009). These interactions are quite frequent in signaling pathways as they provide a mechanism for the cell to quickly respond to extracellular stimuli. They also function in the secretory pathways in eukaryotes by controlling the transport-competent proteins (Nyfeler *et al.*, 2005). Our review mainly focuses on transient complexes. First, we introduce types of protein–protein interactions (PPIs) emphasizing the transient ones. We then describe some key points and case studies about transient protein and drug interactions. Next, we elaborate on the structural characterization of transient and non-obligate complexes, based on the geometrical and physicochemical properties of their interfaces. Kinetic models of transient PPIs (TPPIs) are explained in the proceeding section. We then continue with the introduction of experimental techniques that are used for detecting TPPIs in combination with the computational methods which aim to classify transient and/or non-obligate complexes among other types of PPIs. Furthermore, we list the databases and servers which are utilized to identify and predict TPPIs, and then conclude our review by stating the key points about transient complexes.

Types of protein–protein interactions

Both metabolic and regulatory networks are driven by PPIs; however, different types of complexes with specific functions are observed: large macromolecular complexes, such as the ribosome, are highly stable and permanent whereas dynamic and transient interactions are key components in signaling and regulatory networks (Bhattacharyya *et al.*, 2006; Stein *et al.*, 2009a; Bashor *et al.*, 2010). Protein–protein interactions can be classified based on their composition, affinity and life time (Nooren and Thornton, 2003a; Park *et al.*, 2009) as: (i) homo- and hetero-oligomeric complexes, (ii) non-obligate and obligate complexes (Fig. 1) and (iii) transient and permanent complexes (Fig. 1), respectively.

Homo-oligomeric and hetero-oligomeric complexes

These groups of complexes are differentiated based on their compositions such that if a PPI occurs between identical chains, it is said to form a homo-oligomer whereas if the PPI takes place among non-identical chains then it forms a hetero-oligomer complex. Homo-oligomers are symmetric and provide a good scaffold for stable macromolecules. For example, a chaperonin protein is formed by seven GroEL proteins associating as a homo-heptamer to form a cylinder and seven GroES proteins cap one side of this cylinder (Braig *et al.*, 1994). The cylindrical region is an example of a homo-oligomer, whereas the GroEL/GroES complex is a supramolecule of hetero-oligomers. The stability of hetero-oligomers can vary and form a basis to gather different proteins that cooperate in a single macromolecule. For example, α/β tubulins form a stable dimer and these dimers form long protofilaments, which are constituents of microtubules (Lowe *et al.*, 2001).

Obligate and non-obligate complexes

The key point for differentiation between these two groups is affinity. If the constituents (protomers, monomers) of a complex are unstable on their own *in vivo* then this is an obligate interaction whereas the components of non-obligate interactions can exist independently. As an obligate complex

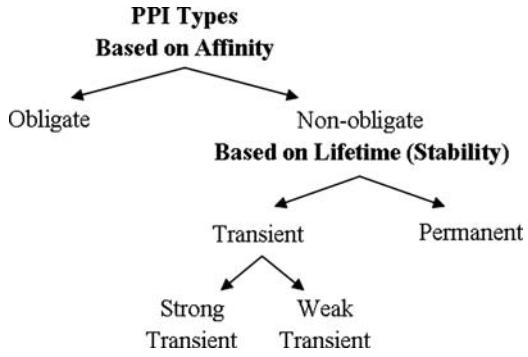

Fig. 1. Relation of protein–protein interaction types based on affinity and stability. Non-obligate interactions are transient but there are some examples of permanent non-obligate interactions such as enzyme–inhibitor interactions (e.g. thrombin–rhodnin inhibitor interaction).

example, Ku proteins, which are involved in DNA repair, are shown to bind DNA as obligate homodimers (Krishna and Aravind, 2010). On the other hand, signaling protein complexes are good non-obligate interaction examples, due to their transient nature. After contributing to the propagation of a signal, they are dissociated into the stable constituent proteins. For example, H-Ras protein, which is a G protein, has a key role in controlling the cell growth and differentiation signaling pathways. It interchangeably forms non-obligate complexes with guanosine triphosphatase (GTPase) activating proteins (GAPs) (acceleration of GDP-bound state of H-Ras—switch OFF) and guanine nucleotide-exchange factors (acceleration of GTP-bound state of H-Ras—switch ON), when the cell is resting and when activated in response to stimuli, respectively (Vetter and Wittinghofer, 2001).

Transient and permanent complexes

These groups of interaction types are discriminated based on the lifetime (or stability) of the complex. Permanent interactions are usually very stable and irreversible (e.g. IL-5 cytokine dimer (PDB ID: 3b5k)) (Nooren and Thornton, 2003b; (Fig. 2a). However, the components of the transient interactions associate and dissociate temporarily *in vivo* (Mintseris and Weng, 2003; Nooren and Thornton, 2003a; Nooren and Thornton, 2003b; Block *et al.*, 2006; Janin *et al.*, 2008; Levy and Pereira-Leal, 2008). The α/β tubulin dimer is an example of an obligate/permanent complex, whereas the dimers of α/β dimers are transient and non-obligatory providing the dynamic nature to microtubules in cell division, cargo transportation and cytoskeleton (Hyams and Lloyd, 1993). Non-obligate interactions are predominantly transient (Janin *et al.*, 2008), with a few examples of permanent (Fig. 1), but obligate interactions are usually permanent in nature (Nooren and Thornton, 2003a). It should be noted that permanent and obligate terms are used interchangeably in the literature.

Transient complexes, depending on their functional roles in the cell, have a wide range of affinities and lifetimes and hence can be further classified as strong and weak (Nooren and Thornton, 2003a; Nyfeler *et al.*, 2005) (see Fig. 2b and c) based on the stability of their oligomeric equilibrium. The strong transient interactions (e.g. heterotrimeric G protein (PDB ID: 1got)) shift equilibrium of association/dissociation under certain disturbances (Nooren and Thornton, 2003a). G proteins, which are crucial in signaling pathways, are examples for strong transient interactions. These membrane-bound proteins are transient in nature and they get activated by G-protein-coupled-receptors (GPCRs) in order to activate a target in the plasma membrane causing a cascade of other signaling events (Alberts *et al.*, 2009). G proteins are composed of three subunits, namely α, β and γ, and are inactive when α subunit is bound to GDP. When α subunit is activated by a ligand-bound GPCR, it loses the affinity for GDP and exchanges it for a molecule of GTP so that the detachment of GTP-bound α subunit from the βγ complex is triggered (see Fig. 2b). However, the weak transient interactions (such as dimers of abalone sperm lysin (PDB ID: 2lyn)) are broken and formed continuously (Nooren and Thornton, 2003a) (see Fig. 2c). The lysin dimer is used to make a hole in the vitelline envelope (VE), which is protective, so that the sperm swims and fuses with the egg (Kresge *et al.*, 2000). For this process to take place, lysin is released from the sperm and binds to the VE receptor lysin (VERL) to dissolve the VE. It was found that the lysin monomer is active during the binding process to VERL, whereas it is observed to be a dimer when it contacts the egg (Kresge *et al.*, 2000).

Domain–domain and domain–peptide complexes

Protein–protein interactions can also be classified based on their folds as domain–domain and domain–peptide interactions (Aloy and Russell, 2006) (see Fig. 2). The complexes belonging to the latter group have mostly transient natures as they are formed by the recognition of a globular domain, a short linear motif (LM) and the small interface on which the interaction takes place (e.g. SH3 domain of tyrosine kinase Fyn—proline-rich peptide (PDB ID: 1fyn) (Stein *et al.*, 2009a; London *et al.*, 2010). These domain–peptide interactions are also called transient peptide-mediated interactions. Indeed, special interaction domains (such as PDZ, SH2, SH3, WW, etc.) provide an elegant mechanism in signaling by making use of transient interactions. These modular interaction domains usually recognize and bind specific motifs of peptides (either at the termini or disordered regions of partner proteins). These are like ready-to-bind interaction domains as they do not undergo large conformational changes on binding and are frequently used. For example, in homo sapiens, there are 223 SH3, 234 PDZ and 91 WW domains (Bhattacharyya *et al.*, 2006). These domains can be used to assemble constituent proteins into large complexes, bringing together different combinations of catalytic domains with regulatory domains. Each complex with a different combination of domains will then have a different function leading to a different signal in the cell.

Transient protein and drug interactions

Cellular processes, such as the cell cycle, which are involved in disease-related pathways, are regulated via transient interactions. Hence, understanding the details of TPPIs by using a systematic wide range approach may enlighten the discovery and development of inhibitors, which can serve as therapeutics for such diseases (Rudolph, 2007; Kar *et al.*, 2010; Ozbabacan *et al.*, 2010). Until recently, TPPI inhibitors were not widely studied as they were considered to bind targets with low specificity, low efficiency, low amount and were difficult to screen and analyze (Rudolph, 2007; Ohlson, 2008). However, experimental and computational advances

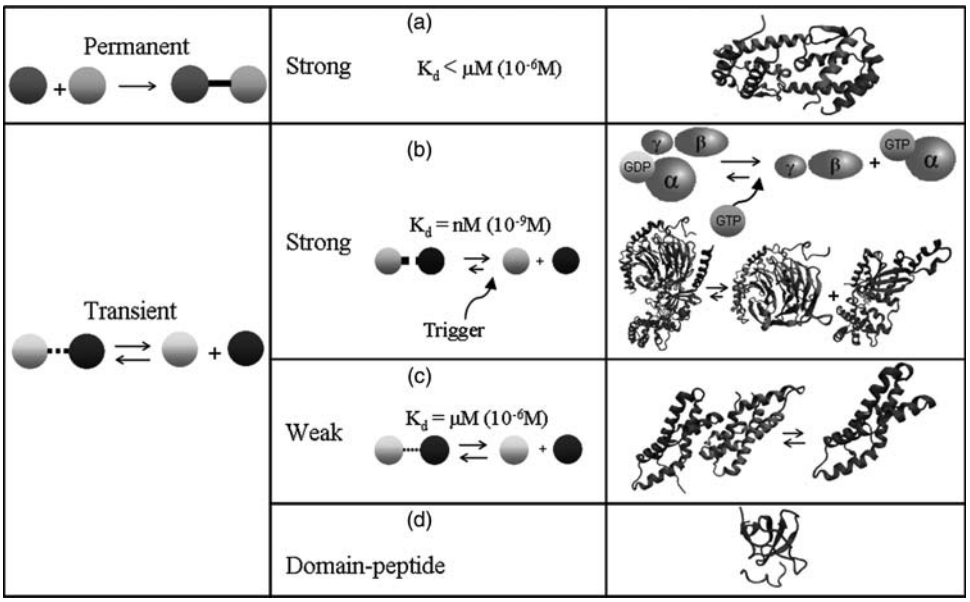

Fig. 2. Classification of protein–protein interaction types based on stability and fold. The mechanisms of association and dissociation processes are shown for strong and weak transient protein–protein complexes, along with the structures of example cases. **(a)** *Permanent protein–protein interaction*: Components are stable only in complex form, e.g. IL-5 cytokine dimer (PDB ID: 3b5k). **(b)** *Strong transient protein–protein interaction*: association/dissociation takes place under certain triggers such as chemical modification, conformational change and colocalization; dissociation constant (K_d) is in nanomolar range, e.g. Heterotrimeric G protein (PDB ID: 1got), in which Gα is in complex with guanosine diphosphate (GDP) and interacts transiently with Gβγ, dissociates into Gα (PDB ID: 3ffb) and Gβγ (PDB ID: 1tbg) subunits upon guanosine triphosphate (GTP) (PDB ID: 3ffb) binding. **(c)** *Weak transient protein–protein interaction*: Complexes are broken and formed continuously and K_d is in micromolar range, e.g. Red abalone lysin dimer (PDB ID: 2lyn) dissociates into Red abalone lysin monomer (PDB ID: 2lis). **(d)** *Domain–peptide interaction*: a globular domain recognizes a short linear motif, e.g. SH3 domain of tyrosine kinase Fyn in complex with a proline-rich peptide (PDB ID: 1fyn).

have significantly improved the knowledge of protein interactions and the inhibitors against them (Rudolph, 2007). Some specific examples of experimental techniques, which use a wide range of proteomic methods, are yeast two-hybrid screens (Uetz *et al.*, 2000), systematic RNA interference (Kamath *et al.*, 2001), mass spectrometry (Ho *et al.*, 2002) and the intracellular localization of proteins with fluorescence markers (Lippincott-Schwartz and Patterson, 2003).

Transient interactions might be important in drug mechanisms in two ways: the drugs that (i) act on TPPIs and that (ii) act transiently on their multiple targets (Rudolph, 2007; Ohlson, 2008). A cancer-related example for the former type of drugs is nutlins (Vassilev *et al.*, 2004). Vassilev *et al.* identified these small and selective inhibitors, which uniquely target the interaction between murine double minute 2 (MDM2) and tumor suppressor p53 in order to stabilize p53, and they developed a novel strategy for cancer therapy. When the interaction between MDM2 and p53 takes place transiently; p53 changes its conformation so that it cannot bind DNA (Wawrzynow *et al.*, 2007) and become degraded. Nutlins inhibit this transient interaction by competitively binding to the site on MDM2, which is the binding site for p53, so that p53 accumulates and becomes activated (Jiang *et al.*, 2007). Another example of a drug which targets TPPIs is brefeldin A (BFA) (Robineau *et al.*, 2000). This inhibitor uncompetitively attacks macromolecular complexes when they are in action; in other words, it attacks when the complex is in a transition state being structurally and energetically unbalanced, so that its hotspots which are targets for drug binding are exposed (Pommier and Cherfils, 2005) (see Fig. 3a). This type of drug mechanism is called 'interfacial inhibition' (Renault *et al.*, 2003). Colchicine is another candidate interfacial inhibitor as it stabilizes the dimer of α-tubulin and β-tubulin by acting on their interface and hence blocking the polymerization (Ravelli *et al.*, 2004) (see Fig. 3b).

A new hypothesis suggests that the multi-target drug approach, which aims at several targets simultaneously, can maximize the efficiency of a drug (Frantz, 2005; Morphy and Rankovic, 2005; Hopkins *et al.*, 2006; Morphy and Rankovic, 2007). The reason to consider multi-target drugs, which may transiently bind to their targets, is the fact that diseases such as cancer, depression, inflammatory and cardiovascular diseases are caused not by a single molecular defect but by dysfunctions that are combined in a complex manner (Ohlson, 2008). These drugs are advantageous in the treatment of neurodegenerative diseases and cancer as they minimize adverse effects, such as cell degeneration, by weakly binding to the receptors and not totally blocking them (Ohlson, 2008). For example, the multi-target anti-cancer agent Gleevec® is found to be promising in leukemia treatment (Frantz, 2005). Furthermore, multi-target drugs are also represented as a form of combinatorial therapy and are more frequently used for the treatment of diseases such as AIDS, cancer and atherosclerosis (Huang, 2002; Borisy *et al.*, 2003; Kaelin, 2004). A few examples of transiently binding drugs are alcohol (ethanol) (Siggins *et al.*, 2005), non-steroidal anti-inflammatory drugs (NSAIDs) namely aspirin, naproxen and ibuprofen (Cryer and Feldman, 1998), weak enzyme inhibitors such as valproic acid and butyric acid (Ohlson, 2008), salicylate and metformin (Csermely *et al.*, 2005). The transient nature of alcohol stems from the fact that it binds to different receptors with low affinity and NSAIDs bind receptors and enzymes with an affinity higher than 1 μM. Memantine is another example of transiently acting drugs and is effective against the group of neurodegenerative

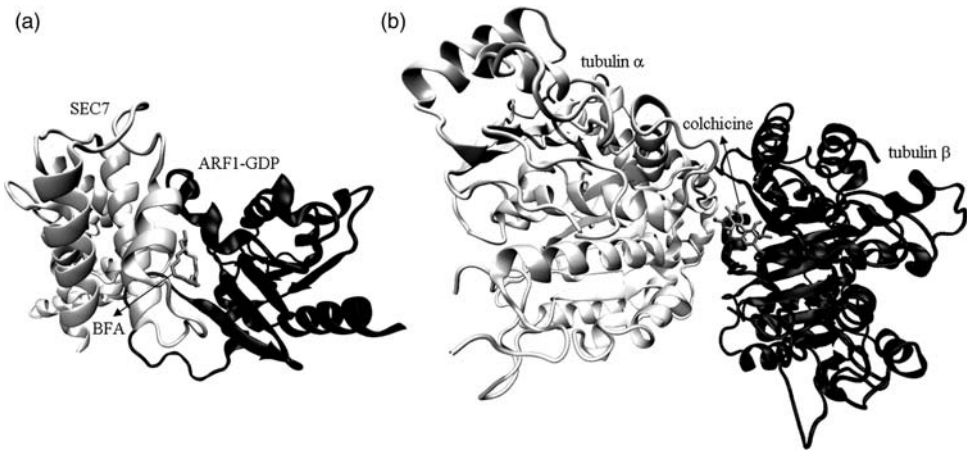

Fig. 3. Examples of drugs acting on transient protein–protein interactions. **(a)** Structure of ARF1-GDP bound to SEC7 domain complexed with Brefeldin A (BFA) (PDB ID: 1re0). **(b)** Structure of tubulin α and tubulin β dimer complexed with colchicine (PDB ID: 1sa0)

diseases called dementia, the most common form of which is Alzheimer's disease (Rogawski, 2000; Lipton, 2004).

Structural characterization of transient protein–protein complexes

A PPI takes place through an interface formed by the interacting pair of proteins and the interface consists of the residues in contact that belong to the chains of proteins on each side. The binding interfaces have been subject to many structural analyses (Jones and Thornton, 1996; Keskin *et al.*, 2008; Tuncbag *et al.*, 2008), as they have key roles not only in the comparison of interaction types but also in the prediction of new PPIs.

With the ever-increasing information on binary PPIs, we know that some proteins bind to 10s or even 100s of other proteins, acting as hubs and these interactions are transient by nature. These hub proteins should have evolved to balance between their specificity and promiscuity (Humphris and Kortemme, 2007; Cukuroglu *et al.*, 2010). Three-dimensional (3D) structures and interfaces of hub proteins with many different partners combined with available kinetic data will surely enlighten our understanding how transient complexes achieve their high specificity and how enzymes bind their inhibitors with high selectivity (e.g. protease-inhibitor complexes) (Meenan *et al.*, 2010). For the structural comparison of permanent and transient interactions, geometrical and physicochemical properties at the interfaces can be considered (Jones and Thornton, 1996). These properties are the change in accessible surface area (ΔASA; as a contact area measure) and planarity as size and shape properties; gap volume index as a measure for complementarity; polarity (Nooren and Thornton, 2003a, Nooren and Thornton, 2003b; Jones and Thornton, 1996), hydrophobicity and mean number of hydrogen bonds; the number of discontinuous segments in the interface for measuring segmentation and the portions of secondary structures; and the extent of conformation change on binding.

In transient domain–peptide interactions, usually the globular recognition domain of a protein (50–150 residues long) interacts with a linear extended peptide (common consensus motif of 3–10 residues) (Stein and Aloy, 2008). The consensus motifs characterizing the peptides are found in loops or unstructured (disordered) regions of proteins. The interfaces of domain–peptide complexes (200–500 Å^2) are smaller than the domain–domain complexes (~2000 Å^2) (Chakrabarti and Janin, 2002; London *et al.*, 2010), and therefore they are characterized as transient (Stein and Aloy, 2008). Usually, the contact area of transient interactions is smaller (than 1500 Å^2) compared with the permanent complexes, which have larger and more twisted interfaces (with contact areas ranging from 1500 to 10000 Å^2) (Nooren and Thornton, 2003a; Nooren and Thornton, 2003b; De *et al.*, 2005; Block *et al.*, 2006; Zhu *et al.*, 2006; Levy and Pereira-Leal, 2008; Park *et al.*, 2009). The limited surface area of the transient (recognition) complex interface is due to the intrinsic physical requirement of the components to fold independently and to exist in solution without aggregating (Mintseris and Weng, 2003). When planarity is considered, the interfaces of heterocomplexes (which may be both transient and permanent) are found to be more planar with respect to homodimers (permanent complexes) (Jones and Thornton, 1996). In terms of polarity as the interfaces of transient complexes resemble the exterior surface of a protein, they have more polar and charged groups (Jones and Thornton, 1997; Lo Conte *et al.*, 1999; Nooren and Thornton, 2003a; Ansari and Helms, 2005). Since permanent complexes are stable, their interfaces are hydrophobic similar to the interior of an average globular protein (Lo Conte *et al.*, 1999; Nooren and Thornton, 2003a; Block *et al.*, 2006; Park *et al.*, 2009). On the other hand, as the components of transient complexes should be stable on their own, their interfaces are less hydrophobic (Jones and Thornton, 1996) and also consist of solvent-exposed amino acids (Tsai *et al.*, 1997). This is intuitive, since transient complexes need to be soluble when dissociated (Block *et al.*, 2006). In terms of residue propensities, obligate (permanent) interactions embody hydrophobic residues such as Leu, Ala, while non-obligate (usually transient) interactions include polar residues such as Ser and Gly (Park *et al.*, 2009). Additionally, both types of interactions include charged residues such as Glu, Asp, Lys (Park *et al.*, 2009) and Arg (De *et al.*, 2005; Park *et al.*, 2009). Mainly non-polar residues such as Ile and Met are observed on the center of obligatory interfaces whereas non-polar residues such as Leu and Val and aromatic residues such as Tyr are included in the core of non-obligatory interfaces (De *et al.*,

Table I. Structural and kinetic characterization of types of protein–protein complexes

	Transient\non-obligate	Permanent\obligate
Interface contact area ΔASA ($Å^2$)	<1500	1500–10000
Secondary structures	Helix and turns	Helix and β-sheet
Interface polarity	High	Low
Conformational changes upon binding	Low	High
Residue propensity	Polar, charged	Hydrophobic, charged
Shape and electrostatic complementarity	High	High
Equilibrium dissociation constant (K_d)	$>10^{-6}$ M (micromolar, μM)	$<10^{-6}$ M (micromolar, μM)

2005). These non-polar residues supply the required strength and specificity to these non-obligate interfaces (De *et al.*, 2005). Although the findings of Park *et al.* (2009) and De *et al.* (2005) agree that the centers of both interface types are mainly non-polar, their findings about residue propensities of non-obligate interactions are contradictory. Trp residue is seen in both types of interfaces both at the center and periphery, with a larger propensity in non-obligatory interfaces (De *et al.*, 2005). This is explained by the fact that Trp is favored as an interaction hotspot (Samanta and Chakrabarti, 2001; Keskin *et al.*, 2005; Ma and Nussinov, 2007).

When the interfaces of obligatory and non-obligatory complexes are compared based on their secondary structures, irregular secondary structure regions such as turns are more frequently involved in non-obligatory interfaces because these regions provide the flexibility required for dissociation under certain conditions (De *et al.*, 2005). Additionally, β-sheet formation is not observed in non-obligatory interfaces because it provides stability (De *et al.*, 2005). For both types of complexes, interactions between two helices are observed. Hence, types of interactions can be characterized by the involvement of regular secondary structures (De *et al.*, 2005; Keskin and Nussinov, 2005).

Salt-bridges and hydrogen bonds take place more often for stabilization of transient complexes (Lo Conte *et al.*, 1999) whereas the association of permanent PPIs may sometimes occur through covalent disulphide bridges between the interacting components (De *et al.*, 2005). Both obligate and transient complexes are close-packed and complementary from the geometric and electrostatic standpoint (Lawrence and Colman, 1993; De *et al.*, 2005; Jones and Thornton, 1995; McCoy *et al.*, 1997; Lo Conte *et al.*, 1999). If an interface belongs to a permanent complex and the conformational rearrangements occurring upon binding take place more often; then the interface size and hydrophobicity are stated to be, respectively, higher (Lo Conte *et al.*, 1999; Nooren and Thornton, 2003a; Nooren and Thornton, 2003b; Levy *et al.*, 2005). Table I summarizes the main points discussed about the structural and kinetic comparison of protein–protein complex types.

The structural interface properties are strongly related with the evolution of complexes. From the evolutionary standpoint, TPPI interfaces are more affected by their local environments or physiological conditions than permanent ones, since they associate or dissociate at least once during their cellular process (Cho *et al.*, 2006). When the interface residues of transient and obligate complexes are compared evolutionarily, their conservations are considerably different. Also, the interface residues of obligate complexes are more robustly dependent on each other than those of transient complexes (Mintseris and Weng, 2005). This is because under high pressure, the coevolution of obligate complexes with their interacting partners enables their interface residues to evolve slowly and at similar rates. Conversely, transient interactions require the fast adaptation of mutations at the interface of the interacting partner causing the correlated mutations to be undetectable (Mintseris and Weng, 2005).

Another evolutionary aspect for comparison of complexes is specificity. The transient interactions between recognition domains and peptides are known to be highly specific *in vivo*, even though the interface consists of a few residues. Also it was found that the binding specificity is dependent on the single-point mutations in arbitrary residues of the consensus motif of a peptide. These facts imply that the transient peptide-mediated interactions evolve in order to maximize specificity (Stein *et al.*, 2009a).

Disordered regions of transient protein–protein complexes

Proteins having disordered regions, which are also called intrinsically unfolded proteins (IUPs) or intrinsically disordered proteins (IDPs), evolve in higher organisms due to their advantageous regulatory strategies. Mutations in these proteins are also directly related to important pathologies in complex organisms such as cancer (Iakoucheva *et al.*, 2002; Kim *et al.*, 2008; Stein *et al.*, 2009a; Uversky *et al.*, 2009). Furthermore, disordered regions of complexes may affect the degree of motion between domains, cover binding sites, act as the targets of post-translational modifications (PTMs), and enable transient binding of different binding partners (Mittag *et al.*, 2010).

Mittag *et al.* (2010) listed some biophysical and evolutionary advantages of IUPs. They mentioned the 'polyelectrostatic' effect that supplies multiple charges on disordered proteins to affect binding affinity through long-range electrostatic effects causing a 'net charge' or 'mean field'. Plasticity and malleability of proteins were stated to expand with the disorder that helped the binding of the same protein sequences with several binding partners. Additionally, since disordered regions show higher rates of mutation they may facilitate restriction of protein size, inhibit molecular crowding inside cells and also limit cell size.

In TPPIs, disordered regions (Dunker *et al.*, 2005; Ekman *et al.*, 2006; Singh *et al.*, 2007) and high content of repeating domains (Bjorklund *et al.*, 2006; Ekman *et al.*, 2006) are important. These properties maintain TPPIs' large surfaces for flexible binding, letting them contain many interactions *in vivo* (Bjorklund *et al.*, 2008). In order to test this hypothesis and describe multi-partnered proteins, hub proteins were classified into two groups as sociable (transient hub proteins) and non-sociable proteins (Higurashi *et al.*, 2008). Since it is now widely accepted that hub proteins tend to have many disordered regions, which will eventually form the binding interfaces (Liu *et al.*, 2002; Wright and Dyson, 1999), the main difference between those two groups is the abundance of disordered regions. However, the most distinctive feature of sociable proteins was found to be the overall structural

flexibility of the proteins, not the disordered regions (Higurashi *et al.*, 2008).

Another relevance of disorder and transient interactions is that transient interactions are likely to be mediated by linear motifs, which are short sequence patterns related to a particular function and generally fall into locally disordered regions. Linear motifs are often found in signaling pathways as consensus sites of PTM or recognition elements in transient complexes, and are identified by local flexibility (Neduva and Russell, 2005; Fuxreiter *et al.*, 2007).

Specific PTMs can take place, in order for the binding sites of the recognition domains to be formed in transient peptide-mediated interactions. The dynamic regulation of complex cellular processes is achieved by these often reversible PTMs (Perkins *et al.*, 2010) (while our work was under revision, this concurrent review which covers other aspects of TPPIs, also appeared) and their fast attachment and removal kinetics. Post-translational modifications of amino acids can occur in many ways such as: addition of simple chemical groups (phosphorylation, methylation and hydroxylation) and the attachment of small proteins (sumoylation and ubiquitylation) (Stein *et al.*, 2009a).

Kinetics of transient protein–protein interactions

Protein complexes have different affinities depending on their functions (Wallis *et al.*, 1995). If the assembly and disassembly of proteins take place in seconds, it is a dynamic process (Levy and Pereira-Leal, 2008). Many complexes in the cell do not last for a long time and dissociate frequently; hence, constituents of these complexes are in equilibrium with the complex form during this dynamic association/dissociation process. The binding affinity of a complex, or in other words the strength of a PPI, may be measured with the equilibrium dissociation constant (K_d). Equilibrium dissociation constant K_d is defined as the ratio between rate constant of the complex dissociation reaction (off rate: k_{off} or k_d) and that of the association reaction (on rate: k_{on} or k_a) (see Scheme 1); and is used to describe the strength of a PPI (Creighton, 1993; Phizicky and Fields, 1995). K_d has the dimension of a concentration and is expressed by molar per liter (also noted as M). As its name suggests, if dissociation constant is high, the reaction tends to proceed in the reverse direction; i.e. the complex tends to dissociate and has a low affinity with a low ratio of bound to free form. Thus, the equilibrium dissociation constants of the transient complexes may be observed in the range of millimolar (10^{-3} M) to micromolar (10^{-6} M), as their constituents associate and dissociate rapidly, whereas the constants of permanent complexes may be in the range of micromolar to femtomolar (10^{-15} M) (Wallis *et al.*, 1995). The knowledge of strong PPIs ($K_d < 10^{-6}$ M) improved considerably over time while transient and/or weak PPIs, especially the ones with $K_d > 10^{-4}$ M, are still poorly understood, although they are known to have essential functions in various cellular processes (Vaynberg *et al.*, 2005). Determination of the high-resolution 3D structures of transient interactions would enlighten the molecular knowledge of specificity and binding mechanisms of weak PPIs. In addition to the two-state kinetics (see Scheme 1) for protein association introduced above; three-state and four-state association kinetics are also observed for transient interactions.

Three-state kinetics assumes that the process of association goes through an intermediate state (A*B) called transient complex (Zhou *et al.*, 1997; Alsallaq and Zhou, 2008) or sometimes encounter complex (Gabdoulline and Wade, 1997). The transient complex can either dissociate into its components or form the final native state (AB) (see Scheme 2, Alsallaq and Zhou, 2008). The overall rate constant of association (k_a, $M^{-1}s^{-1}$) is dependent on diffusion-controlled rate constant (k_D), dissociation rate constant of the transient complex (k_{-D}) and conformational rearrangement rate constant (k_c) (Scheme 2). The observed protein association rate constants (k_a's) range from 10^3 to $10^9\,M^{-1}s^{-1}$ (Schreiber *et al.*, 2009). Four-state kinetic models are also used to describe the association of a protein complex (Schreiber, 2002). According to this model (Scheme 3), A and B proteins form an unstable encounter complex (AB*) initially and then it evolves into the intermediate (AB**), which will finally form the final complex (AB). The association rates are usually between 10^5 and $10^6\,M^{-1}s^{-1}$ but sometimes it can exceed $10^9\,M^{-1}s^{-1}$ for interactions in which speed is important (Schreiber, 2002).

$$A + B \underset{k_d}{\overset{k_a}{\rightleftarrows}} AB; \quad K_d \equiv \frac{k_d}{k_a} \tag{1}$$

$$A + B \underset{k_{-D}}{\overset{k_D}{\rightleftarrows}} A*B \xrightarrow{k_c} AB; \quad k_a = \frac{k_D k_c}{k_{-D} + k_c} \tag{2}$$

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} AB^* \underset{k_{-2}}{\overset{k2}{\rightleftarrows}} AB^{**} \underset{k_{-3}}{\overset{k3}{\rightleftarrows}} AB \tag{3}$$

Non-obligate and obligate complexes have been designated alternative names based on their functions. For instance, the key roles of non-obligate proteins in signaling pathways cause them to also be called recognition complexes whereas obligate ones are known as folding complexes for being formed as a result of protein biosynthesis (Block *et al.*, 2006). Obligate and non-obligate complexes are also designated as two-state and three-state complexes, respectively (Tsai *et al.*, 1998; Xu *et al.*, 1998; Mintseris and Weng, 2003). Obligate complexes are called two-state as they function in processes where inseparable binding and folding take place, whereas the name of the latter type comes from the fact that components of three-state (non-obligate) complexes fold independently and then associate with each other (Mintseris and Weng, 2003). Some authors also address the obligate and non-obligate complexes, based on their specifities, as cognate and noncognate, respectively (Wallis *et al.*, 1995; Meenan *et al.*, 2010). Cognate complexes are regarded as high-affinity complexes, whose interacting partners are specific to each other such as the complexes of colicin endonucleases (DNases) with immunity (Im) proteins (Meenan *et al.*, 2010). Colicins are protein antibiotics that target *Escherichia coli* cells and the Group E colicins, through their DNase domain, degrade the bacterial genome (Meenan *et al.*, 2010). The organism can protect itself only by a specific Im protein, which binds to the incoming colicin. Colicin DNase–Im complexes are of special importance in terms of specificity, as their equilibrium dissociation constant values cover a wide range of stabilities (Meenan *et al.*, 2010). Only a specific Im can bind to colicin, which is an incoming cytotoxin, in order to protect the organism. On the other hand, a non-cognate complex is composed of weakly associated

Table II. Comparison of experimental methods for detection of transient protein–protein interactions

Method	Advantages	Disadvantages
Y2H	Scalable, eligible of analyzing many interactions	Low covering of detected interactions between different experiments (Bader and Chant, 2006). High false-positive (false negative) rate
Mammalian two hybrid	Highly complementary to Y2H with regard to the subset of interactions they are able to detect (Lievens *et al.*, 2009)	Much less responsive to high-throughput analyses than yeast technologies such as Y2H (Lievens *et al.*, 2009)
NMR	Investigating weak PPIs at atomic levels (Vaynberg and Qin, 2006)	Abundance of data obtained from a system
TRNOESY	No isotope labeling is a necessity (Vaynberg and Qin, 2006)	The target protein masses have to be large enough (Vaynberg and Qin, 2006)
TAP-TAG	High probability of detecting actual protein partners quantitatively *in vivo* (Collins and Choudhary, 2008)	Transient interactions are believed to be lost throughout long purification time (Collins and Choudhary, 2008)
FRET	Following transient interactions precisely with respect to time and high resolution in single cells (Phizicky *et al.*, 2003)	When it fails, there will be uncertainty about the cause since it may be due to false location, free fluorophores or proteins' distance to each other
YFP-PCA	Direct visualization of protein–protein interactions in their normal compartmental environment of living cells (Nyfeler *et al.*, 2005)	Increased solubility of YFP fragments (Nyfeler *et al.*, 2005)
BIFC	A simple and sensitive method due to the stability of the reconstituted YFP complexes (Hu *et al.*, 2002; Ohad *et al.*, 2007)	Detection of non-specific interactions when expression levels of the split YFP fragments are high (Hu *et al.*, 2002; Ohad *et al.*, 2007)
SPR	Examining quantity of formed complex in the presence of free material. Does not require a washing process before quantization (Rich and Myszka, 2007)	Not efficient for high-throughput assays. Not very reliable for analyzing small molecules

proteins, and can be observed between DNase and Im couples that are not specific to each other (Meenan *et al.*, 2010). The binding affinity of such complexes might be much weaker than the cognate complexes. For example, the authors found that the binding affinity of the cognate colicin E9 endonuclease (E9 DNase)–immunity protein 9 (Im9) interaction is seven orders of magnitude higher than the non-cognate E9 DNase–Im2 complex, yet they observed similar types of hotspots and conserved interfacial water interactions in both complexes. These hotspot interactions are so favorable that they can even tolerate some other unfavorable interactions leading to selectivity for different partner proteins.

Experimental detection of transient protein–protein interactions

Identification of TPPIs, which occur instantaneously, is technically challenging because they produce an insufficient amount of complexes and cannot be easily recognized *in vitro* or *in vivo* by traditional approaches. Classical methods used in biotechnology may lead to picking the most robust complexes, while weakly bound and transient complexes might be ignored. Identification and analysis of TPPIs require sensitive and high-resolution experimental techniques (Sali *et al.*, 2003). Some of the high-resolution analysis techniques, which detect TPPIs directly, are listed below. Advantages and disadvantages of the experimental methods described in this section are summarized in Table II.

Nuclear magnetic resonance (NMR) spectroscopy is one of the most useful tools for investigating weak protein–target interactions at physiological conditions (Qin *et al.*, 2001; Walters *et al.*, 2001; Zuiderweg, 2002; Gao *et al.*, 2004), and it is effective for investigating the weak PPIs at atomic levels (Vaynberg and Qin, 2006). Nuclear magnetic resonance effect is observed when magnetic nuclei take in and diffuse electromagnetic energy in a magnetic field. It was first described by Isidor Rabi in 1938 (Rabi *et al.*, 1992). Paramagnetic relaxation enhancement (PRE) is one of the NMR approaches. It maintains a method for directly investigating the presence and the nature of low population, transient intermediates under equilibrium conditions (Iwahara and Clore, 2006). Data on complexes in the fast exchange regime, obtained from PRE, supplies useful information about intermediates. These observations reveal both the structural features and the presence of intermediate states. Another NMR procedure is 2D transferred nuclear Overhauser effect spectroscopy (TRNOESY), which is also a quick assay for identifying weak PPIs. Since no isotope labeling is a necessity, this method is known to be economic. One handicap of the system is that the target protein masses have to be large enough (Vaynberg and Qin, 2006). Kobayashi *et al.* used this method to study the interaction in solution between minichaperone GroEL (193–335) and a synthetic peptide (Rho) (Kobayashi *et al.*, 1999).

Disulfide trapping is an effective method of obtaining further structural information about weak interactions in the guidance of NMR docking. This approach was used on a non-cognate complex between the colicin E9 endonuclease (E9 DNase) and immunity protein 2 (Im2) (Meenan *et al.*, 2010).

Another suitable method for the identification and analysis of transient interactions is fluorescence resonance energy transfer (FRET). This method detects the interactions based on physical distance, in which energy can be transferred from an excited molecular flurophore (the donor) to another fluorophore (the acceptor) via intermolecular long-range dipole–dipole coupling (Sekar and Periasamy, 2003). Fluorescence resonance energy transfer enables experimentalists to follow the transient interactions with precise measurements with respect to time and high resolution in single cells. Measurements yielded by this method supply a non-invasive procedure to visualize the spatiotemporal dynamics of interactions between protein partners *in vivo* (Sullivan and Kay, 1999; Phizicky *et al.*, 2003). As a tool that enables detection of inter- and intramolecular interactions of fluorescent proteins, FRET has a major role in modern fluorescence microscopy (Gertler *et al.*, 2005) (Fig. 4a).

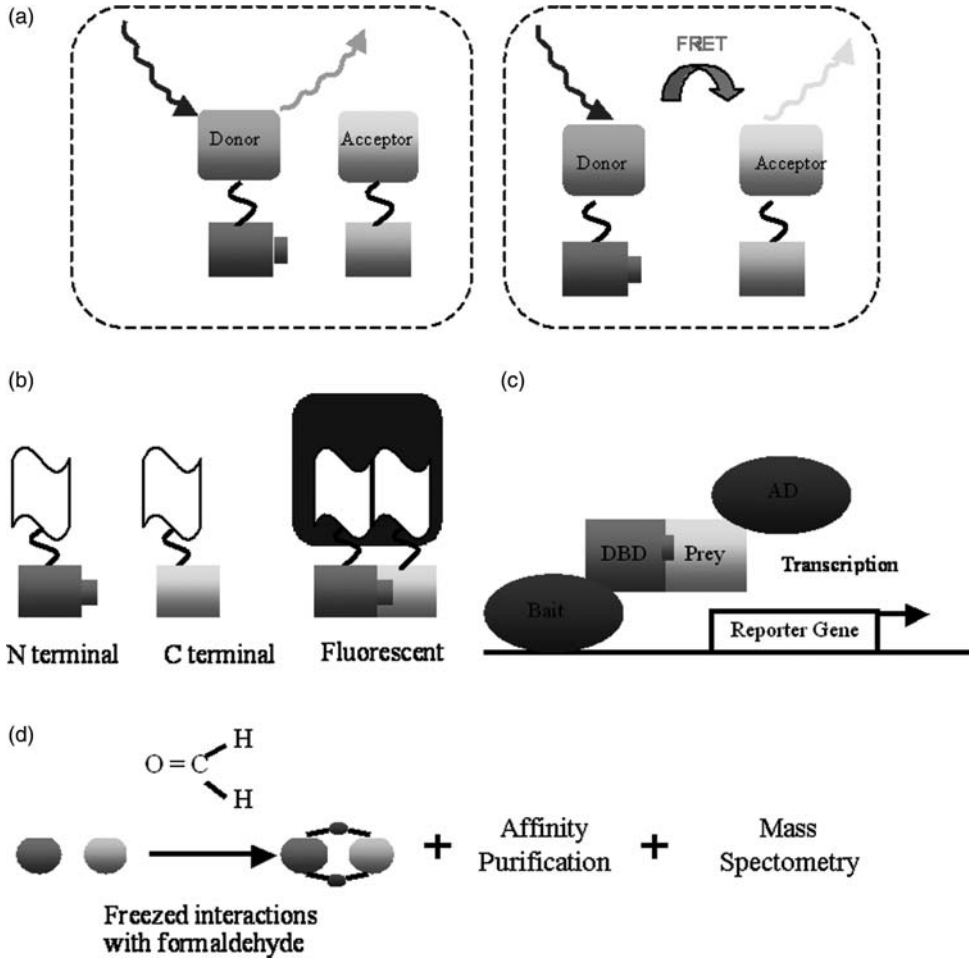

Fig. 4. Experimental methods used for the detection of transient protein-protein interactions. **(a)** *Fluorescence resonance energy transfer:* this procedure, in which energy is transferred from an excited molecular flurophore (the donor) to another fluorophore (the acceptor) via intermolecular long-range dipole–dipole coupling, detects the interactions based on the physical distance (Sekar and Periasamy, 2003). **(b)** *Bimolecular fluorescence complementation:* in the BFC method, N and C terminals of YFP/GFP are fused to two distinct proteins; in the case of meeting again, they fluorophore. **(c)** *Yeast two-hybrid method:* this method is based on the molecular dissection of transcription activators. Since a particular structural contact is not required between DNA binding and transcriptional activation domains, the physical connection among them can be substituted with a non-covalent interaction settled by interacting proteins (Seraphin, 2002). **(d)** *Chemical cross-linking:* with the help of formaldehyde, chemical cross-linking is used to stabilize interactions through covalent bond constructions in weak and/or transient native cells or tissues, during the purification processes.

The bimolecular fluorescence complementation (BIFC) method can be used as an alternative method to FRET. Green fluorescent protein (GFP) family is used for interaction trapping. It is shown to be appropriate primarily for protein interaction identification in bacteria (Ghosh *et al.*, 2000) and in mammalian cells (Hu and Kerppola, 2003). The use of the BIFC method for direct visualization of weak intracellular protein interactions is emphasized by Morell *et al.* (2007). The working principle of this method is the binding of split yellow fluorescent protein/green fluorescent protein (YFP/GFP) variants in order to construct a functional fluorophore (Ohad *et al.*, 2007). Morell *et al.* (2007) used SH3 domain with natural and designed binding partners as a test case (Fig. 4b).

Nyfeler *et al.* (2005) applied a YFP-based protein fragment complementation assay (PCA) to secretory pathway of living cells for identifying PPIs. Secretory pathway is a challenging focus due to the transient nature of the interactions it contains; such as the interaction between proteins of the endoplasmic reticulum quality control machinery and their substrates or the interaction between cargo and cargo receptors. Detection of low-affinity interactions was achieved by fixing the complex by the reconstituted YFP. Yellow fluorescent protein PCA could visualize weak, transient protein interactions that may escape interest by coimmunoprecipitation and chemical cross-linking.

Surface plasmon resonance (SPR) which makes use of an optical event, was first practiced in biology in 1983 (Liedberg *et al.*, 1983). It is a useful method for screening transiently binding proteins in real time. The main advantage of SPR biosensor technology is its ability to examine the quantity of formed complex in the presence of free material. It also does not require a washing process before quantitation. That is why transient interactions can be characterized with this method (Rich and Myszka, 2007). In addition, SPR maintains a great variety of information about molecular interactions, such as specificity and the kinetic details on binding such as strength, affinity and the rate of association. On the other hand, this mechanism is not very reliable for analyzing small molecules (Pattnaik, 2005). Further, one should keep in mind that SPR data may give different rate constants from the ones obtained in solution (Schreiber *et al.*, 2009).

Apart from the methods explained above, some of the experimental techniques that detect TPPIs indirectly are presented below.

The yeast two-hybrid (Y2H) system, which is being performed in many laboratories, is one of the most popular methods for detecting weak PPIs. This system was originally developed in 1989 by Fields and Song (1989) and the premise behind it was based on the molecular dissection of transcription activators (Fig. 4c). Since a particular chemical bond is not required between DNA binding and transcriptional activation domains, the physical connection between them can be substituted with a non-covalent interaction settled by interacting proteins (Seraphin, 2002). Because the genetic reporter gene strategy concludes with remarkable signal amplification (Estojak *et al.*, 1995), transient and weak interactions are usually identified by this procedure (Berggard *et al.*, 2007). The major handicap in forming interactome maps with the data obtained from different Y2H experiments is the low overlap of detected interactions between different experiments (Bader and Chant, 2006). Performing consecutive Y2H screening is proposed as a solution for handling this problem (Venkatesan *et al.*, 2009; Vinayagam *et al.*, 2010). This procedure captures transient interactions that cannot be detected in a single Y2H experiment and increases the overlap between different datasets. In order to detect 90% of all Y2H detectable interactions, at least six repeated Y2H screens are needed (Venkatesan *et al.*, 2009). Vinayagam *et al.* (2010) found that singletons, which are interactions found in a single experiment, are transient in nature. Despite being widely used, the Y2H system is known to produce a high number of false positives (and false negatives). Consequently, there is a validation requirement for the interactions discovered by Y2H.

The mammalian two hybrid system, which is also capable of detecting transient and weak interactions, is known as a complementary method to the Y2H system. This method was first defined by Rossi *et al.* (1997), who monitored PPIs in intact eukaryotic cells by β–galactosidase complementation. While the sensitivity of the method is comparable to Y2H, these two methods are mostly complementary to each other in terms of the interactions they detect (Lievens *et al.*, 2009). In a mammalian host, mammalian proteins are likely to preserve their native conformation, which permits PPIs to be recorded as a function of time, space and physiological context (Fu and Liang, 2002; Lievens *et al.*, 2009).

Affinity chromatography-based methods, which are commonly used for separating biochemical mixtures in many laboratories, may not be sufficient for transient protein interactions since they tend to bias toward high-affinity interactions and slow kinetics of dissociation, particularly when stringent rinsing processes are performed. The fact that the dilute buffers used in laboratories and intracellular environments are very different from each other may be one explanation for this bias. In fact, PPIs *in vivo* take place in a highly condensed macromolecular mixture. This intracellular protein concentration affects the diffusion rate of molecules and cause competition for water. As a result, the connection of two proteins may have much higher affinity in a loaded environment like the milieu inside a cell than a buffer (Berggard *et al.*, 2007).

Parallel to those processes, transient interactions are believed to be lost throughout the long purification time needed for the Tandem Affinity Purification method, whereas single-step purifications could preserve them relatively. Labeling of low abundance transient interactions may be handicapped because of the reduced purity of single-step methods (Collins and Choudhary, 2008). That is why mass spectrometry analysis is unable to characterize transient complexes when coupled with affinity purification strategies. Nevertheless, with the help of cross-linking *in vivo*, transient interactions may be frozen through covalent-bond formation before affinity purification. Formaldehyde is commonly used as a cross-linker for this purpose (Orlando *et al.*, 1997; Ethier *et al.*, 2006). The quantitative analysis of the tandem-affinity-purified *in vivo* cross-linked protein complexes (QTAX) method (Guerrero *et al.*, 2006; Guerrero *et al.*, 2008) is an example of integrated mass spectrometry based procedures. With the help of formaldehyde, chemical cross-linking is used to stabilize interactions through covalent bond constructions in weak and/or transient native cells or tissues, during purification processes (Fig. 4d). Other than QTAX, a number of techniques have recently been developed for differentiation of static and dynamic interactions based on the tandem affinity purification method. Two of these methods are time course-purification after mixing-SILAC (Tc-PAM-SILAC) and mixing after purification-SILAC (MAPSILAC) (Mousson *et al.*, 2008; Wang and Huang, 2008).

The phage display is an alternative method for detecting PPIs. It is a display method of polypeptides or proteins via fusion to phage coat proteins, and it uses bacteriophages for linking proteins with the genetic information encoding them (Smith, 1985). Like the tandem-affinity-purification method, this technique is not suitable for identifying transient protein interactions.

Computational studies on transient protein–protein interactions

The previous section listed some of the experimental methods to detect transient and non-obligate PPIs. Recently, many researchers have also focused on diverse computational methods for the prediction and classification of PPI types. Computational methods combined with the experimental detection results improve the overall understanding about TPPIs.

Distinguishing obligate complexes from non-obligate ones

Structural interface properties of complexes are widely used to distinguish between types such as obligate and non-obligate complexes. The following computational studies are based on differentiating between types of complexes by using their interfacial properties, and hence they require initial knowledge of 3D structures.

In 2003, Nooren and Thornton (2003b) compared weak transient homodimers with obligate heterodimers via structurally analyzing their interface properties. Results suggested that although identification of interaction types based solely on interface properties is difficult, the ASA and polarity of the interface are critical parameters in distinguishing transient complexes from the more stable and obligate ones.

De *et al.* (2005) focused on classification of proteins into two sets as obligatory and non-obligatory (or transient) complexes. They statistically analyzed the structural descriptors of interfaces, such as area and polarity, on the chain level for recognition of those two classes. Statistical tests on different features pointed out that only some of the features are

significant. So, they benefited from the cumulative effect of known specifications—like obligatory interfaces having larger interface areas and being non-polar, and involving secondary structural elements across the interface—in the classification step. They also found out that the stability of a complex can be judged by the information about the interface properties.

A machine learning-based classification (support vector machine) by Zhu *et al.* (2006) differentiated biological interactions from crystal packing contacts, and differentiated obligate interactions from non-obligate ones. Their algorithm, called NOXclass, uses six attributes, namely interface area, ratio of interface area to protein surface area, amino acid composition of the interface, correlation between amino acid compositions of interface and protein surface, interface shape complementarity and conservation of the interface. Using the leave one-out cross-validation procedure, NOXclass obtains 91.8% accurate classifications. Additionally, NOXclass empowers the prediction and evaluation of protein quaternary structures and gives clues about the features of PPIs when experimental data are unavailable.

Atomic contact vectors (ACV), introduced by Mintseris and Weng (2003), are representations of atomic contacts in interfaces and are used to compare the interfaces of transient recognition and permanent oligomeric complexes. After finding all protein–protein interfaces available in PDB (Berman *et al.*, 2000), the researchers were able to distinguish these two types of complexes with 91% accuracy by using ACVs. Later, they also compared these two classes of complexes on an evolutionary basis (Mintseris and Weng, 2005). They concluded that the transient complexes are members of larger families compared with obligate ones and the number of paralogs per represented species is higher in transient complexes.

Distinguishing permanent complexes from transient ones

Distinguishing via structural interface properties. Similarly, structural interface properties are used in the literature for evaluating whether a complex is transient or permanent. Gunasekaran *et al.* (2004) analyzed the interface and surface properties, such as ASA, residue composition and polarity, of several types of complexes, consisting of natively unstructured proteins, ribosomal proteins, two-state and three-state complexes and crystal dimers. Starting from known 3D structures, they centered their research on whether ordered and disordered monomers are dissimilar in their structural properties, existing in their complexed form. Results revealed that two-state (permanent) complexes coincide with disordered proteins whereas the three-state (transient) ones coincide with crystal-packing dimers. Ordered proteins' per residue interface and surface areas were found to be considerably smaller than the disordered proteins. Making use of this fact, a simple scale that evaluates whether a protein in its complex form can exist as a stable monomer, was provided (Gunasekaran *et al.*, 2004). They also introduced a scheme to classify whether the proteins in complexes are ordered (stable) or disordered when separated from their partners.

Using Minsteris and Weng's (Mintseris and Weng, 2003) structurally known interface dataset, Block *et al.* (2006) classified transient and permanent complexes. The physical and chemical properties of those complexes' interfaces were represented at the atomic level. By using machine learning algorithms as a means of classification, their work obtained 93.6% accuracy. Results highlighted the significance of the contact area as a discriminating property between permanent and transient complexes. They also acquired 76% accurate predictions by just focusing on the sizes of the interfaces.

Additionally, depending on residue–residue preferences and sequence properties, such as amino acid composition, an information theory-based method for differentiating between different types of interactions, including transient and permanent interactions, was developed by Ofran and Rost (2003). Without any prior knowledge of protein complex structures and based only on the amino acid composition, they were able to statistically predict the class of an interface correctly in 63–100% of the cases.

Recently, Park *et al.* (2009) used pattern discovery of the interaction sites for the classification of a set of structurally known complexes, which represent four different interaction types. These interaction types are obligate permanent interactions consisting of homo or hetero obligomers and non-obligate transient interactions consisting of enzyme inhibitor or non-enzyme inhibitor. They also reinforced the prediction of PPI types using association rule-based classification (ARBC). Their results revealed that prediction efficiency of classification models may considerably benefit from the selective capability of association rules. Additionally, this work showed that structural domain information and secondary structure content may improve classification accuracy.

Distinguishing via expression data. Some computational methods consider expression levels as a classification criterion for complexes and do not require a solved protein structure. An example for comparing transient and permanent complexes was the work of Jansen *et al.* (2002). They studied the relationship between mRNA expression levels and the type of PPIs via computationally clustering and inter-relating the expression levels of different data sources for yeast. Two different types of expression measurements were used: absolute expression levels in vegetative yeast cells (SAGE or gene chip experiments) and ratio-type expression data from microarray experiments. They pointed out a strong correlation between expression levels and permanent protein complexes whereas transient complexes were found to be weakly correlated with expression data. Based on the fact that permanent complexes are known to be coexpressed, whereas the coexpression of transient complexes is lower, transient interactions are harder to identify with coexpression data.

Tirosh and Barkai (2005) introduced a verification tool for PPIs, depending on the coexpression of orthologs of interacting partners. They proved that the expression data from multiple organisms can lead to increased confidence of hypothetical PPIs by analyzing coexpression of orthologs of the presumed interacting partners. Making use of this concept, coexpression of orthologs was shown to be especially useful for identification of transient interactions.

Another method of detecting candidate PPIs based on gene expression data was published by Zanivan *et al.* (2007). In this work they claimed that interacting pairs belonging to a multi-protein complex could be more easily detected because correlations in expression data show much higher signal-to-noise ratio when multiple correlations are considered at the same time. An additional outcome of this study was the combination of the standard Pearson-based

Table III. Databases and servers for identification and prediction of transient protein–protein interactions

Server/database	Explanation	Web site
ELM	Largest linear motif server	http://elm.eu.org/
PROSITE	Linear motif server	http://expasy.org/prosite/
SCANSITE	Linear motif server	http://scansite.mit.edu/
ADAN	Predictive analysis of modular domains settled by linear motifs	http://adan.crg.es/
SPIDER	PPI site recognition	http://sppider.cchmc.org/
NOXCLASS	Classification of obligate, non-obligate and crystal packing interactions	http://noxclass.bioinf.mpiinf.mpg.de/
PRISM	Protein interaction prediction by structural matching	http://prism.ccbb.ku.edu.tr/
3DID	Structural explanation of how protein interactions arise in molecular level	http://3did.irbbarcelona.org/
Minimotif Miner	Queries for the presence of short functional motifs	http://mnm.engr.uconn.edu/
QuasiMotiFinder	Identification of signatures and signature-like patterns in protein sequences	http://quasimotifinder.tau.ac.il/
AutoMotifServer	Identification of post-translational modification sites in proteins	http://ams2.bioinfo.pl/
SIRW	Search protein/nucleotide databases with a sequence motif	http://sirw.embl.de/
DILIMOT	Finding linear motifs, in a set of protein	http://dilimot.embl.de/
SLiMFinder	Linear Motif Server	http://bioinformatics.ucd.ie/shields/software/slimfinder/

method, which is capable of permanent interaction determination, with the analyses of synchronous peaks of expression. This combined method was more suitable for TPPI predictions since the expression peaks method can detect the functionally important changes in gene expression although the level of change is small.

Overall, machine learning approaches are widely used in differentiating transient interactions from permanent ones. Among interface attributes, ASA seems to be the major discriminator. Other than interface properties, microarray expression data were also shown to be useful in identification of transient interactions.

Databases and servers for identification and prediction of TPPIs

As explained in the previous sections, capturing TPPIs experimentally or computationally is a tedious and hard assignment. Thus, our comprehension of these short-time interactions is limited. As an attempt for advancing our understanding about the issue, knowledge obtained by predictions and experiments have been collected in databases and servers, some of which are listed below. A list of these databases and servers for identification and prediction of PPIs are given in Table III.

Eukaryotic LM (ELM) database (Puntervoll *et al.*, 2003) which is available at http://elm.eu.org/, contains many peptide-mediated transient interactions of eukaryotic proteins. Motifs in this database are usually between 4 and 11 residues long, which might not be noticed in high-throughput experiments (Pawson and Linding, 2005) but ELM supplies a literature-originated cluster of motifs and their interaction partners (Stein and Aloy, 2008). Although it is the largest linear motif database, there are many other linear motif servers such as PROSITE (Bairoch, 1993) and Scansite (Obenauer *et al.*, 2003).

The Adan database (Encinar *et al.*, 2009) is built for predictive analysis of modular domains settled by linear motifs. It combines different modular protein domains (SH2, SH3, PDZ and WW) and has a subset name Prediadan, which maintains position-specific scoring matrices for PPIs and predictions of optimum ligands and candidate binding partners. The Adan database is accessible at http://adan-embl.ibmc.umh.es/ or http://adan.crg.es/.

Sppider (Porollo and Meller, 2007) is a tool for PPI site recognition (available at http://sppider.cchmc.org/), which combines enhanced relative solvent accessibility predictions with high-detailed structural data. Porollo and Meller used the NT86 dataset of transient complexes (Nooren and Thornton, 2003b), to test the success of their tool. Sppider can be classified as a useful prediction tool for transient complexes, with 74% accuracy.

The PPI classification tool of Park *et al.* (2009), as introduced in the previous chapter, is based on association rules and has a web application available at http://bioinfo.ssu.ac.kr/~shpark/picasso/. This tool identifies different PPI types such as obligate, permanent and non-obligate transient interactions.

NOXclass (Zhu *et al.*, 2006) server, which was also described in the previous section, is used for classifying obligate, non-obligate and crystal packing interactions. This server is available at http://noxclass.bioinf.mpiinf.mpg.de/.

Tuncbag *et al.* (2008) published a dataset (PRINT) consisting of 8205 interface clusters, each standing for a separate interface structure. In this dataset, which is available from http://prism.ccbb.ku.edu.tr/, there are 14 501 obligate and 2709 non-obligate PPIs, classified with NOXclass (Zhu *et al.*, 2006) filter. Those interfaces are classified into three types as type I clusters; similar interface architectures, similar global folds, type II clusters; similar interface architectures, dissimilar global folds, type III clusters; one side structurally similar interfaces (Keskin *et al.*, 2004). Type III cluster consists of multipartnered and transient interfaces (Fig. 5).

PRISM is a novel algorithm for predicting PPIs based on structure conservation in protein interfaces (Aytuna *et al.*, 2005; Ogmen *et al.*, 2005). PRISM bases its predictions on template interface datasets. When a subset of the non-obligate template dataset, containing 158 TPPIs is used, one can get predictions for other potential transient interactions. This method, which is now being studied by many others, was the first to make PPI predictions based on the structural similarity of the binding sites. It is publicly available from http://prism.ccbb.ku.edu.tr/.

3did is a database which uses structural data in order to explain how protein interactions arise at the molecular levels (Stein *et al.*, 2005) and is available at http://3did.irbbarcelona.org/. It contains both globular domain interactions and a hand-curated set of transient peptide-mediated interactions (Stein *et al.*, 2009b). If 3did is queried with a specific

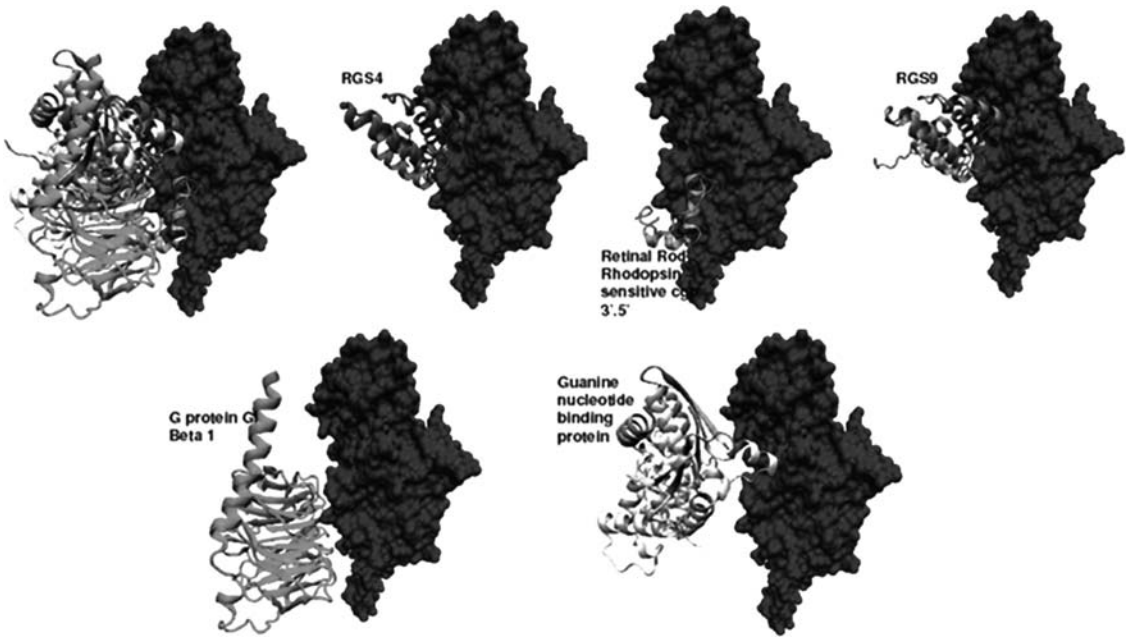

Fig. 5. Multipartnered and transient interfaces of a protein. Actin complexes are examples of type III interfaces (Tuncbag *et al.*, 2008), which are transient and multipartnered. The individual interactions of the partners of Guanine nucleotide binding protein are not possible simultaneously. Although the binding regions of some partners are not overlapping, the rest of the partner proteins interpenetrate each other.

domain/motif, a protein sequence, PDB codes or gene ontology (GO) terms via the web-based tool; all resulting domains or peptides which connect with the specified domain with a known 3D structure are displayed. Furthermore, those interaction partners are shown in an interactive network which reveals facts about the type of interaction (domain–domain or peptide-mediated), whether they are intra- or intermolecular and which enables the user to choose the depth and a color scheme based on molecular function, biological process or cellular compartment as described by GO.

Although they are not specific to transient interactions, there are many online PPI databases that can be used to mine transient interactions. String (Snel *et al.*, 2000; Jensen *et al.*, 2009) is one of those. It merges data from various sources including experimental results, computational predictions and the literature. Both physical and functional interactions of proteins can be explored via String. iRefIndex is another resource which is an index of protein interactions, gathered from numerous principal PPI databases such as BIND (Bader *et al.*, 2003; Alfarano *et al.*, 2005), BioGRID (Stark *et al.*, 2006), DIP (Salwinski *et al.*, 2004), HPRD (Peri *et al.*, 2003; Mishra *et al.*, 2006), IntAct (Aranda *et al.*, 2010), MINT (Chatr-aryamontri *et al.*, 2007), MPact (Guldener *et al.*, 2006), MPPI (Pagel *et al.*, 2005) and OPHID (Brown and Jurisica, 2005). Both String and iRefIndex provide a non-repetitious index of interactors for a specific protein (Razick *et al.*, 2008).

The Critical Assessment of Predicted Interactions (CAPRI) (Janin *et al.*, 2003) is a community wide experiment performed multiple times a year. It is assessed to answer a specific question: 'Given the 3D structure of elements of a complex how certain would be the predictive models?'. Although it is not specific to transient interaction prediction; mining this resource may be suitable for enhancing our knowledge on transient interactions.

Concluding remarks

The components of transient complexes associate and dissociate rapidly while transiently interacting with each other to function dynamically in crucial regulatory and signaling pathways. The identification and analysis of these complexes have become more manageable with the emerging sensitive and high-resolution experimental techniques accompanied by the high-throughput computational methods.

However, to better understand and eventually design a TPPI, a molecular perspective of the protein–protein interfaces is necessary. Obtaining high-quality structures could be one way of achieving this goal as structural information can help us in several ways. For example, in a recent study, using the NMR-based docking and disulfide trapping strategy, the co-crystal structure of a transient E9 Dnase–Im2 complex was obtained. This helped to show, on a molecular level, how specificity can evolve within this family where the members have a wide range of affinity and specificities. Their results indicate that the core of the interface provides high stability to the complexes whereas the surrounding residues determine the specificity of the complexes. Furthermore, the advances in NMR and X-ray crystallography techniques can lead to more and higher quality structures of transient complexes. As the coverage of these techniques increase, they can provide a good template to understand and design new transient complexes. An example for such advanced techniques is, PRISM, which uses available transient interactions as a template set and searches structural and evolutionary similarities between the template set and the target proteins to be predicted.

Acknowledgements

This work was supported by the Scientific and Technological Research Council of Turkey Fellowship (Grant Numbers 109T343 and 109E207 to S.E.A.O. and H.B.E.) and Turkish Academy of Sciences (TUBA).

References

Alberts,B., Bray,D., Hopkin,K., Johnson,A., Lewis,J., Raff,M., Roberts,K. and Walter,P. (2009) *Essential Cell Biology*. 3 edn, Garland Science, New York.

Alfarano,C., Andrade,C.E., Anthony,K. *et al*. (2005) *Nucleic Acids Res.*, **33**, D418–D424.

Aloy,P. and Russell,R.B. (2006) *Nat. Rev. Mol. Cell Biol.*, **7**, 188–197.

Alsallaq,R. and Zhou,H.X. (2008) *Proteins*, **71**, 320–335.

Ansari,S. and Helms,V. (2005) *Proteins*, **61**, 344–355.

Aranda,B., Achuthan,P., Alam-Faruque,Y. *et al*. (2010) *Nucleic Acids Res.* **38**, D525–D531.

Aytuna,A.S., Gursoy,A. and Keskin,O. (2005) *Bioinformatics*, **21**, 2850–2855.

Bader,G.D., Betel,D. and Hogue,C.W. (2003) *Nucleic Acids Res.*, **31**, 248–250.

Bader,J.S. and Chant,J. (2006) *Science*, **311**, 187–188.

Bairoch,A. (1993) *Nucleic Acids Res.*, **21**, 3097–3103.

Bashor,C.J., Horwitz,A.A., Peisajovich,S.G. and Lim,W.A. (2010) *Annu. Rev. Biophys.*, **39**, 515–537.

Berggard,T., Linse,S. and James,P. (2007) *Proteomics*, **7**, 2833–2842.

Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) *Nucleic Acids Res.*, **28**, 235–242.

Bhattacharyya,R.P., Remenyi,A., Yeh,B.J. and Lim,W.A. (2006) *Annu. Rev. Biochem.*, **75**, 655–680.

Bjorklund,A.K., Ekman,D. and Elofsson,A. (2006) *PLoS Comput. Biol.*, **2**, e114.

Bjorklund,A.K., Light,S., Hedin,L. and Elofsson,A. (2008) *Proteomics*, **8**, 4657–4667.

Block,P., Paern,J., Hullermeier,E., Sanschagrin,P., Sotriffer,C.A. and Klebe,G. (2006) *Proteins*, **65**, 607–622.

Borisy,A.A., Elliott,P.J., Hurst,N.W. *et al*. (2003) *Proc. Natl Acad. Sci. USA*, **100**, 7977–7982.

Braig,K., Otwinowski,Z., Hegde,R., Boisvert,D.C., Joachimiak,A., Horwich,A.L. and Sigler,P.B. (1994) *Nature*, **371**, 578–586.

Brown,K.R. and Jurisica,I. (2005) *Bioinformatics*, **21**, 2076–2082.

Chakrabarti,P. and Janin,J. (2002) *Proteins*, **47**, 334–343.

Chatr-aryamontri,A., Ceol,A., Palazzi,L.M., Nardelli,G., Schneider,M.V., Castagnoli,L. and Cesareni,G. (2007) *Nucleic Acids Res.*, **35**, D572–574.

Cho,K.I., Lee,K., Lee,K.H., Kim,D. and Lee,D. (2006) *Proteins*, **65**, 593–606.

Collins,M.O. and Choudhary,J.S. (2008) *Curr. Opin. Biotechnol.*, **19**, 324–330.

Creighton,T.E. (1993) *Proteins: Structures and Molecular Properties*. 2nd edn, W.H. Freeman, New York.

Cryer,B. and Feldman,M. (1998) *Am. J. Med.*, **104**, 413–421.

Csermely,P., Agoston,V. and Pongor,S. (2005) *Trends Pharmacol. Sci.*, **26**, 178–182.

Cukuroglu,E., Gursoy,A. and Keskin,O. (2010) *Ann. Biomed. Eng.*, **38**, 2068–2078.

De,S., Krishnadev,O., Srinivasan,N. and Rekha,N. (2005) *BMC Struct. Biol.*, **5**, 15.

Dunker,A.K., Cortese,M.S., Romero,P., Iakoucheva,L.M. and Uversky,V.N. (2005) *FEBS J.*, **272**, 5129–5148.

Ekman,D., Light,S., Bjorklund,A.K. and Elofsson,A. (2006) *Genome Biol.*, **7**, R45.

Encinar,J.A., Fernandez-Ballester,G., Sanchez,I.E., Hurtado-Gomez,E., Stricher,F., Beltrao,P. and Serrano,L. (2009) *Bioinformatics*, **25**, 2418–2424.

Estojak,J., Brent,R. and Golemis,E.A. (1995) *Mol. Cell Biol.*, **15**, 5820–5829.

Ethier,M., Lambert,J.P., Vasilescu,J. and Figeys,D. (2006) *Anal. Chim. Acta.*, **564**, 10–18.

Fields,S. and Song,O. (1989) *Nature*, **340**, 245–246.

Frantz,S. (2005) *Nature*, **437**, 942–943.

Fu,L. and Liang,J.J. (2002) *J. Biol. Chem.*, **277**, 4255–4260.

Fuxreiter,M., Tompa,P. and Simon,I. (2007) *Bioinformatics*, **23**, 950–956.

Gabdoulline,R.R. and Wade,R.C. (1997) *Biophys. J.*, **72**, 1917–1929.

Gao,G., Williams,J.G. and Campbell,S.L. (2004) *Methods Mol. Biol.*, **261**, 79–92.

Gertler,A., Biener,E., Ramanujan,K.V., Djiane,J. and Herman,B. (2005) *J. Dairy Res.*, **72 Spec No.**, 14–19.

Ghosh,I., Hamilton,A.D. and Regan,L. (2000) *J. Am. Chem. Soc.*, **122**, 5658–5659.

Guerrero,C., Milenkovic,T., Przulj,N., Kaiser,P. and Huang,L. (2008) *Proc. Natl Acad. Sci. USA*, **105**, 13333–13338.

Guerrero,C., Tagwerker,C., Kaiser,P. and Huang,L. (2006) *Mol. Cell Proteomics*, **5**, 366–378.

Guldener,U., Munsterkotter,M., Oesterheld,M., Pagel,P., Ruepp,A., Mewes,H.W. and Stumpflen,V. (2006) *Nucleic Acids Res.*, **34**, D436–441.

Gunasekaran,K., Tsai,C.J. and Nussinov,R. (2004) *J. Mol. Biol.*, **341**, 1327–1341.

Higurashi,M., Ishida,T. and Kinoshita,K. (2008) *Protein Sci.*, **17**, 72–78.

Ho,Y., Gruhler,A., Heilbut,A. *et al*. (2002) *Nature*, **415**, 180–183.

Hopkins,A.L., Mason,J.S. and Overington,J.P. (2006) *Curr. Opin. Struct. Biol.*, **16**, 127–136.

Hu,C.D., Chinenov,Y. and Kerppola,T.K. (2002) *Mol. Cell.*, **9**, 789–798.

Hu,C.D. and Kerppola,T.K. (2003) *Nat. Biotechnol.*, **21**, 539–545.

Huang,S. (2002) *Drug Discov. Today*, **7**, S163–S169.

Humphris,E.L. and Kortemme,T. (2007) *PLoS Comput. Biol.*, **3**, e164.

Hyams,J.S. and Lloyd,C.W. (1993) *Microtubules, Modern Cell Biology*. Wiley-Liss, New York.

Iakoucheva,L.M., Brown,C.J., Lawson,J.D., Obradovic,Z. and Dunker,A.K. (2002) *J. Mol. Biol.*, **323**, 573–584.

Iwahara,J. and Clore,G.M. (2006) *Nature*, **440**, 1227–1230.

Janin,J., Bahadur,R.P. and Chakrabarti,P. (2008) *Q Rev. Biophys.*, **41**, 133–180.

Janin,J., Henrick,K., Moult,J., Eyck,L.T., Sternberg,M.J., Vajda,S., Vakser,I. and Wodak,S.J. (2003) *Proteins*, **52**, 2–9.

Jansen,R., Greenbaum,D. and Gerstein,M. (2002) *Genome Res.*, **12**, 37–46.

Jensen,L.J., Kuhn,M., Stark,M. *et al*. (2009) *Nucleic Acids Res.*, **37**, D412–416.

Jiang,M., Pabla,N., Murphy,R.F., Yang,T., Yin,X.M., Degenhardt,K., White,E. and Dong,Z. (2007) *J. Biol. Chem.*, **282**, 2636–2645.

Jones,S. and Thornton,J.M. (1995) *Prog. Biophys. Mol. Biol.*, **63**, 31–65.

Jones,S. and Thornton,J.M. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 13–20.

Jones,S. and Thornton,J.M. (1997) *J. Mol. Biol.*, **272**, 121–132.

Kaelin,W.G., Jr. (2004) *Sci. STKE*, **2004**, pe12.

Kamath,R.S., Martinez-Campos,M., Zipperlen,P., Fraser,A.G. and Ahringer,J. (2001) *Genome Biol.*, **2**, RESEARCH0002.

Kar,G., Keskin,O., Gursoy,A. and Nussinov,R. (2010) *Curr. Opin. Pharmacol.*, **10**, 715–722.

Keskin,O., Gursoy,A., Ma,B. and Nussinov,R. (2008) *Chem. Rev.*, **108**, 1225–1244.

Keskin,O., Ma,B. and Nussinov,R. (2005) *J. Mol. Biol.*, **345**, 1281–1294.

Keskin,O. and Nussinov,R. (2005) *Protein Eng. Des. Sel.*, **18**, 11–24.

Keskin,O., Tsai,C.J., Wolfson,H. and Nussinov,R. (2004) *Protein Sci.*, **13**, 1043–1055.

Kim,P.M., Sboner,A., Xia,Y. and Gerstein,M. (2008) *Mol. Syst. Biol.*, **4**, 179.

Kobayashi,N., Freund,S.M., Chatellier,J., Zahn,R. and Fersht,A.R. (1999) *J. Mol. Biol.*, **292**, 181–190.

Kresge,N., Vacquier,V.D. and Stout,C.D. (2000) *Acta. Crystallogr. D Biol. Crystallogr.*, **56**, 34–41.

Krishna,S.S. and Aravind,L. (2010) *J. Struct. Biol.*, **172**, 294–299.

Lawrence,M.C. and Colman,P.M. (1993) *J. Mol. Biol.*, **234**, 946–950.

Levy,E.D. and Pereira-Leal,J.B. (2008) *Curr. Opin. Struct. Biol.*, **18**, 349–357.

Levy,Y., Cho,S.S., Onuchic,J.N. and Wolynes,P.G. (2005) *J. Mol. Biol.*, **346**, 1121–1145.

Liedberg,B.O., Nylander,C. and Lundstrom,I. (1983) *Sensors and Actuators*, **4**, 299–304.

Lievens,S., Lemmens,I. and Tavernier,J. (2009) *Trends Biochem. Sci.*, **34**, 579–588.

Lippincott-Schwartz,J. and Patterson,G.H. (2003) *Science*, **300**, 87–91.

Lipton,S.A. (2004) *Nature*, **428**, 473.

Liu,J., Tan,H. and Rost,B. (2002) *J. Mol. Biol.*, **322**, 53–64.

Lo Conte,L., Chothia,C. and Janin,J. (1999) *J. Mol. Biol.*, **285**, 2177–2198.

London,N., Movshovitz-Attias,D. and Schueler-Furman,O. (2010) *Structure*, **18**, 188–199.

Lowe,J., Li,H., Downing,K.H. and Nogales,E. (2001) *J. Mol. Biol.*, **313**, 1045–1057.

Ma,B. and Nussinov,R. (2007) *Curr. Top Med. Chem.*, **7**, 999–1005.

McCoy,A.J., Chandana Epa,V. and Colman,P.M. (1997) *J. Mol. Biol.*, **268**, 570–584.

Meenan,N.A., Sharma,A., Fleishman,S.J., Macdonald,C.J., Morel,B., Boetzel,R., Moore,G.R., Baker,D. and Kleanthous,C. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 10080–10085.

Mintseris,J. and Weng,Z. (2003) *Proteins*, **53**, 629–639.

Mintseris,J. and Weng,Z. (2005) *Proc. Natl Acad. Sci. USA*, **102**, 10930–10935.

Mishra,G.R., Suresh,M., Kumaran,K. *et al*. (2006) *Nucleic Acids Res.*, **34**, D411–414.

Mittag,T., Kay,L.E. and Forman-Kay,J.D. (2010) *J. Mol. Recognit.*, **23**, 105–116.
Morell,M., Espargaro,A., Aviles,F.X. and Ventura,S. (2007) *Proteomics*, **7**, 1023–1036.
Morphy,R. and Rankovic,Z. (2005) *J. Med. Chem.*, **48**, 6523–6543.
Morphy,R. and Rankovic,Z. (2007) *Drug Discov. Today*, **12**, 156–160.
Mousson,F., Kolkman,A., Pijnappel,W.W., Timmers,H.T. and Heck,A.J. (2008) *Mol. Cell Proteomics*, **7**, 845–852.
Neduva,V. and Russell,R.B. (2005) *FEBS Lett.*, **579**, 3342–3345.
Nooren,I.M. and Thornton,J.M. (2003a) *EMBO J.*, **22**, 3486–3492.
Nooren,I.M. and Thornton,J.M. (2003b) *J. Mol. Biol.*, **325**, 991–1018.
Nyfeler,B., Michnick,S.W. and Hauri,H.P. (2005) *Proc. Natl Acad. Sci. USA*, **102**, 6350–6355.
Obenauer,J.C., Cantley,L.C. and Yaffe,M.B. (2003) *Nucleic Acids Res.*, **31**, 3635–3641.
Ofran,Y. and Rost,B. (2003) *J. Mol. Biol.*, **325**, 377–387.
Ogmen,U., Keskin,O., Aytuna,A.S., Nussinov,R. and Gursoy,A. (2005) *Nucleic Acids Res.*, **33**, W331–W336.
Ohad,N., Shichrur,K. and Yalovsky,S. (2007) *Plant Physiol.*, **145**, 1090–1099.
Ohlson,S. (2008) *Drug Discov. Today*, **13**, 433–439.
Orlando,V., Strutt,H. and Paro,R. (1997) *Methods*, **11**, 205–214.
Ozbabacan,S.E., Gursoy,A., Keskin,O. and Nussinov,R. (2010) *Curr. Opin. Drug Discov. Dev.*, **13**, 527–537.
Pagel,P., Kovac,S., Oesterheld,M. *et al.* (2005) *Bioinformatics*, **21**, 832–834.
Park,S.H., Reyes,J.A., Gilbert,D.R., Kim,J.W. and Kim,S. (2009) *BMC Bioinform.*, **10**, 36.
Pattnaik,P. (2005) *Appl. Biochem. Biotechnol.*, **126**, 79–92.
Pawson,T. and Linding,R. (2005) *FEBS Lett.*, **579**, 1808–1814.
Peri,S., Navarro,J.D., Amanchy,R. *et al.* (2003) *Genome Res.*, **13**, 2363–2371.
Perkins,J.R., Diboun,I., Dessailly,B.H., Lees,J.G. and Orengo,C. (2010) *Structure*, **18**, 1233–1243.
Phizicky,E., Bastiaens,P.I., Zhu,H., Snyder,M. and Fields,S. (2003) *Nature*, **422**, 208–215.
Phizicky,E.M. and Fields,S. (1995) *Microbiol. Rev.*, **59**, 94–123.
Pommier,Y. and Cherfils,J. (2005) *Trends Pharmacol. Sci.*, **26**, 138–145.
Porollo,A. and Meller,J. (2007) *Proteins*, **66**, 630–645.
Puntervoll,P., Linding,R., Gemund,C. *et al.* (2003) *Nucleic Acids Res.*, **31**, 3625–3630.
Qin,J., Vinogradova,O. and Gronenborn,A.M. (2001) *Methods Enzymol.*, **339**, 377–389.
Rabi,II., Zacharias,J.R., Millman,S. and Kusch,P. (1992) *J. Magn. Reson. Imaging*, **2**, 131–133.
Ravelli,R.B., Gigant,B., Curmi,P.A., Jourdain,I., Lachkar,S., Sobel,A. and Knossow,M. (2004) *Nature*, **428**, 198–202.
Razick,S., Magklaras,G. and Donaldson,I.M. (2008) *BMC Bioinform.*, **9**, 405.
Renault,L., Guibert,B. and Cherfils,J. (2003) *Nature*, **426**, 525–530.
Rich,R.L. and Myszka,D.G. (2007) *Anal. Biochem.*, **361**, 1–6.
Robineau,S., Chabre,M. and Antonny,B. (2000) *Proc. Natl Acad. Sci. USA*, **97**, 9913–9918.
Rogawski,M.A. (2000) *Amino Acids*, **19**, 133–149.
Rossi,F., Charlton,C.A. and Blau,H.M. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 8405–8410.
Rudolph,J. (2007) *Nat. Rev. Cancer*, **7**, 202–211.
Sali,A., Glaeser,R., Earnest,T. and Baumeister,W. (2003) *Nature*, **422**, 216–225.
Salwinski,L., Miller,C.S., Smith,A.J., Pettit,F.K., Bowie,J.U. and Eisenberg,D. (2004) *Nucleic Acids Res.*, **32**, D449–451.
Samanta,U. and Chakrabarti,P. (2001) *Protein Eng.*, **14**, 7–15.
Schreiber,G. (2002) *Curr. Opin. Struct. Biol.*, **12**, 41–47.
Schreiber,G., Haran,G. and Zhou,H.X. (2009) *Chem. Rev.*, **109**, 839–860.
Sekar,R.B. and Periasamy,A. (2003) *J. Cell Biol.*, **160**, 629–633.
Seraphin,B. (2002) *Adv. Protein Chem.*, **61**, 99–117.
Siggins,G.R., Roberto,M. and Nie,Z. (2005) *Pharmacol. Ther.*, **107**, 80–98.
Singh,G.P., Ganapathi,M. and Dash,D. (2007) *Proteins*, **66**, 761–765.
Smith,G.P. (1985) *Science*, **228**, 1315–1317.
Snel,B., Lehmann,G., Bork,P. and Huynen,M.A. (2000) *Nucleic Acids Res.*, **28**, 3442–3444.
Stark,C., Breitkreutz,B.J., Reguly,T., Boucher,L., Breitkreutz,A. and Tyers,M. (2006) *Nucleic Acids Res.*, **34**, D535–539.
Stein,A. and Aloy,P. (2008) *PLoS One*, **3**, e2524.
Stein,A., Pache,R.A., Bernado,P., Pons,M. and Aloy,P. (2009a) *FEBS J.*, **276**, 5390–5405.
Stein,A., Panjkovich,A. and Aloy,P. (2009b) *Nucleic Acids Res.*, **37**, D300–D304.
Stein,A., Russell,R.B. and Aloy,P. (2005) *Nucleic Acids Res.*, **33**, D413–D417.
Sullivan,K.F. and Kay,S.A. (1999) Cytoskeletal Dynamics in Yeast. In Wilson,L. and Matsudaira,P. (eds.), *Green Fluorescent Proteins*. Academic Press, San Diego.
Tirosh,I. and Barkai,N. (2005) *BMC Bioinform.*, **6**, 40.
Tsai,C.J., Lin,S.L., Wolfson,H.J. and Nussinov,R. (1997) *Protein Sci.*, **6**, 53–64.
Tsai,C.J., Xu,D. and Nussinov,R. (1998) *Fold Des.*, **3**, R71–R80.
Tuncbag,N., Gursoy,A., Guney,E., Nussinov,R. and Keskin,O. (2008) *J. Mol. Biol.*, **381**, 785–802.
Uetz,P., Giot,L., Cagney,G. *et al.* (2000) *Nature*, **403**, 623–627.
Uversky,V.N., Oldfield,C.J., Midic,U., Xie,H., Xue,B., Vucetic,S., Iakoucheva,L.M., Obradovic,Z. and Dunker,A.K. (2009) *BMC Genomics*, **10**(Suppl 1), S7.
Valdar,W.S. and Thornton,J.M. (2001) *Proteins*, **42**, 108–124.
Vassilev,L.T., Vu,B.T., Graves,B. *et al.* (2004) *Science*, **303**, 844–848.
Vaynberg,J., Fukuda,T., Chen,K., Vinogradova,O., Velyvis,A., Tu,Y., Ng,L., Wu,C. and Qin,J. (2005) *Mol. Cell*, **17**, 513–523.
Vaynberg,J. and Qin,J. (2006) *Trends Biotechnol.*, **24**, 22–27.
Venkatesan,K., Rual,J.F., Vazquez,A. *et al.* (2009) *Nat. Methods*, **6**, 83–90.
Vetter,I.R. and Wittinghofer,A. (2001) *Science*, **294**, 1299–1304.
Vinayagam,A., Stelzl,U. and Wanker,E.E. (2010) *Theor. Chem. Acc.*, **125**, 613–619.
Wallis,R., Leung,K.Y., Pommer,A.J., Videler,H., Moore,G.R., James,R. and Kleanthous,C. (1995) *Biochemistry*, **34**, 13751–13759.
Walters,K.J., Ferentz,A.E., Hare,B.J., Hidalgo,P., Jasanoff,A., Matsuo,H. and Wagner,G. (2001) *Methods Enzymol.*, **339**, 238–258.
Wang,X. and Huang,L. (2008) *Mol. Cell Proteomics*, **7**, 46–57.
Wawrzynow,B., Zylicz,A., Wallace,M., Hupp,T. and Zylicz,M. (2007) *J. Biol. Chem.*, **282**, 32603–32612.
Wright,P.E. and Dyson,H.J. (1999) *J. Mol. Biol.*, **293**, 321–331.
Xu,D., Tsai,C.J. and Nussinov,R. (1998) *Protein Sci.*, **7**, 533–544.
Zanivan,S., Cascone,I., Peyron,C., Molineris,I., Marchio,S., Caselle,M. and Bussolino,F. (2007) *Genome Biol.*, **8**, R256.
Zhou,H.X., Wong,K.Y. and Vijayakumar,M. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 12372–12377.
Zhu,H., Domingues,F.S., Sommer,I. and Lengauer,T. (2006) *BMC Bioinform.*, **7**, 27.
Zuiderweg,E.R. (2002) *Biochemistry*, **41**, 1–7.