Local Interaction Signal Analysis Predicts Protein-Protein Binding Affinity

Raffaele Raucci, Elodie Laine, Alessandra Carbone

Local Interaction Signal Analysis predicts protein-protein binding affinity

Raffaele Raucci[1,2], Elodie Laine[1] and Alessandra Carbone[1,3*]

[1] Sorbonne Universités, CNRS, IBPS, Laboratoire de Biologie Computationnelle et Quantitative (LCQB), 4 place Jussieu, 75005 Paris, France

[2] Sorbonne Université, Institut des Sciences du Calcul et des Données (ISCD), 75005 Paris, France

[3] Institut Universitaire de France, Paris 75005, France

April 5, 2018

*Lead contact and corresponding author: alessandra.carbone@lip6.fr

Abstract

Several models estimating the strength of the interaction between proteins in a complex have been proposed. By exploring the geometry of contact distribution at protein-protein interfaces, we provide an improved model of binding energy. Local Interaction Signal Analysis (LISA) is a radial function based on terms describing favorable and non-favorable contacts obtained by Density Functional Theory, the Support-Core-Rim interface residue distribution, non-interacting charged residues and secondary structures contribution. The three-dimensional organisation of the contacts and their contribution on localised hot-sites over the entire interaction surface were numerically evaluated. LISA achieves a correlation of 0.81 (and RMSE of 2.35 ± 0.38 kcal/mol) when tested on 125 complexes for which experimental measurements were realised. LISA's performance is stable for subsets defined by functional composition and extent of conformational changes upon complex formation. A large-scale comparison with 17 other functions demonstrated the power of the geometrical model in the understanding of complex binding.

Software availability: LISA package is freely available at `http://www.lcqb.upmc.fr/LISA/` under the CeCILL licence.

Key words: binding affinity, buried surface area, non-interacting surface, atom-atom contact, non-covalent interaction, favorable contact, non-favorable contact, protein contact, protein-protein complex, protein-protein interaction, electron density, reduced density gradient, secondary structure, protein-protein interface, interface geometry, contact distribution, Density Functional Theory.

Introduction

Proteins communicate by physical interactions. Understanding the way a protein interacts with its partners (Jones & Thornton 1996, Nooren & Thornton 2003, Ubbink 2009, Perkins et al. 2010) and its natural preference to associate to other proteins (Sacquin-Mora et al. 2008, McGuffee & Elcock 2010, Lopes et al. 2013, Laine & Carbone 2016), that is its binding affinity, is of fundamental importance to describe its behaviour, its function, and more generally, its network of protein-protein interactions (PPIs) (Hakes et al. 2008). This means to learn the general principles of protein-protein interface contacts, the favorable signals supporting the affinity of the interaction and the location of favorable contacts at the interface (Chen et al. 2013, Erijman et al. 2014). Since PPIs alterations are often the cause of cell disfunctioning, this understanding is expected to have immediate implications on disease research (Keskin et al. 2005, Aloy & Russell 2006, Beltrao et al. 2007, Kiel et al. 2008, Dell'Orco 2009), drug design (Hartwell et al. 1999, Zhao & Chmielewski 2005, Betzi et al. 2007), computational mutagenesis (Ben-Shimon & Eisenstein 2010) and protein engineering (Kortemme et al. 2004, Sharabi et al. 2011).

Binding affinity defines whether or not complex formation occurs and it is described through the equilibrium dissociation constant K_d, or equivalently the Gibbs free energy ($\Delta G = RT \ln K_d$). Because expensive and time-costly techniques are required to experimentally measure the value K_d, various computational methods aimed at predicting binding affinity have been developed. Among them free energy perturbation and thermodynamic integration techniques are applicable to only a few complexes because of their high computational cost. More efficient approaches exploiting only the three-dimensional structure of the complex have been proposed for the past two decades. They estimate the binding affinity by using force-field potentials, statistical potentials, and docking scores (Horton & Lewis 1992, Moal et al. 2011, Jiang et al. 2002, Ma et al. 2002, Audie & Scarlata 2007, Zeng & Li 2008, Su et al. 2009, Bai et al. 2011, Tian et al. 2012, Zhou et al. 2013). Very recently, a strickingly simple method (Vangone & Bonvin 2015), called Prodigy, based on the counting of atom-atom contacts at the interface and on the charge distribution at the non-interacting surface of the complex, was shown to reach the best correlations with experimental data.

In parallel to the algorithmic development, an important effort was made to establish a faithful benchmark dataset of experimental binding affinity measures that could be used as a reference for a balanced assessment of the methods (Kastritis & Bonvin 2010, Kastritis et al. 2011, Fleishman et al. 2011). The last updated version (Vreven et al. 2015) comprises 179 protein-protein complexes whose structures have been solved at high resolution, as well as those of their unbound components, and for which dissociation constants have been measured by biophysical methods.

Based on the observation that the number of interatomic contacts at the interface gives a crucial contribution to binding affinity (Vangone & Bonvin 2015), here, we investigate whether a more sophisticated description of these contacts, based on quantum chemistry, can further improve accuracy in binding affinity prediction.

Recent developments in quantum chemistry originated from the need of a fast qualitative visualisation of non-covalent interactions in complex molecular systems. Indeed, non-covalent interactions are characterised by low electron density and only slight variations of them, challenging their extraction and characterisation based on the signed electron density and the reduced gradient, two scalar fields derived from quantum mechanical electron density, the key quantity in Density Functional Theory (DFT). A combinatorial algorithm for the automated extraction of these quantities in biological systems of variable size (from interacting nucleobases and simple dimers to protein-ligand, protein-protein and protein-DNA complexes) was proposed (Johnson et al. 2010). This computationally fast algorithm (based on approximate

promolecular density) runs with only the molecular geometry (atomic coordinates) of the complex as input. Most importantly, for a ligand-protein interaction, it allows to localise the many small contributions with a much finer description than the atom-specific description. In fact, van der Waals, dipole-dipole and hydrophobic interactions are not atom-specific and occupy broader regions in space, that can be computed as continuous surfaces, more precisely isosurfaces, rather than close contacts between atom pairs.

In this work, our basic idea is that, to evaluate the binding affinity of a ligand to a protein, one should take into consideration all small non-covalent contributions, and that a spatial analysis of these contributions on isosurfaces could provide a more realistic model of binding affinity and, a deeper understanding of protein-protein interactions. To demonstrate this, we introduce an empirical function for the prediction of binding affinity, LISA (Local Interaction Signal Analysis), based on an original study of the geometrical distribution of atom-atom contacts, favouring the interaction between two proteins, that is guided by the analysis of isosurfaces. LISA is a non-linear combination of several terms computed from four main blocks of analysis. First, quantum mechanical electron density and the two scalar fields derived from it, the reduced gradient (RDG) and the signed electron density, are used in LISA to simultaneously explore a wide range of non-covalent interaction types (van de Waals interactions, hydrogen-bonds, dipole-dipole interactions, steric repulsions, London dispersion (Kollman 1977)) as isosurfaces. Such surfaces allow us to distinguish favorable from non-favorable contacts, and to take into account only specific regions in space that contribute to the protein-protein interaction. Second, the geometry of the interaction site is explicitly considered through its organisation in Support-Core-Rim (SCR) regions (Levy 2010). This allows us to quantify the spread of the signal on isosurfaces with respect to the SCR structure. Third, LISA model considers the effect of the non-interacting surface which was shown to contribute significantly to the binding affinity (Kastritis et al. 2014, Vangone & Bonvin 2015, Marillet et al. 2016). In LISA, we also tested the implication of secondary structures lying at the interface and we could demonstrate that regular structural elements also play some role in the interaction. The ten LISA terms were selected starting from a large set of about 200 features by using objective criteria and a fully automatic selection procedure.

To train and test LISA, we used the subset of 125 complexes of Affinity Benchmark Version 2 (Vreven et al. 2015), for which reliable experimental techniques (Vangone & Bonvin 2015) have been used to measure the binding affinity. It comprises a large variety of complexes with diversified functional activity and reliable experimental measurements of binding affinity. On it, we compared LISA against a large number of tools available and demonstrated that LISA is stable for subsets of Affinity Binding 2 defined by functional composition and extent of conformational changes upon complex formation. LISA outperforms all existing comparable predictor methods (based on no conformational sampling) (Moal et al. 2011, Vangone & Bonvin 2015) and applies to a large variety of datasets of complexes resulting in a very stable behaviour.

LISA model highlights that 'hot-sites' presenting medium/strong concentration of (favorable) atom-atom contacts play a major contribution in binding affinity and that general geometrical principles guide their distribution at interacting surfaces. These findings open the way to further investigations on the geometrical characteristics of those regions.

RESULTS

We present an empirical binding affinity function, LISA, based on a fine modelling of atom-atom contacts. LISA localises the forces at the interface that influence the most the affinity between two proteins and suggests general principles of the distribution of these forces on interaction surfaces. In LISA, the effect

of atom-atom contacts on binding affinity is studied through an analysis of the geometry of the atomic interactions realised at three different levels (local, medium and global) of granularity. First, LISA explores local interferences among co-localized atom-atom contacts to distinguish favorable from non-favorable contacts. Second, it investigates hot-sites of interactions by identifying regions in the protein interfaces with high or medium interaction strength, and high or medium number of contacts. Third, it explores the global distribution of contacts over the entire interaction surface. The combination of the information extracted at the three levels allows LISA to outperform over current binding affinity models. We evaluate LISA on a set of 125 complexes from the structure-based Affinity Benchmark Version 2 (Vreven et al. 2015), which we call AffinityB2_rel, for which experimental measurements of binding affinity have been realised with reliable techniques (Vangone & Bonvin 2015). We show that LISA sharply improves over current predictors.

Modeling the spatial distribution of non-covalent atom-atom contacts

Non-covalent interactions are crucial for describing the interplay of structure and reactivity in large biomolecular systems, and our aim is to define a model describing in the most precise manner general geometric principles that place hot-sites of medium/high strength contacts in specific regions of the interaction surface. Note that here, the term "hot-site" refers to regions (possibly involving several residues, and possibly comprising only a portion of the atom-atom contacts within a residue-residue contact) of the interacting surface, in contrast to the term "hot spot", generally referring to residues that play a particular role in the interaction and/or bringing an important energetic contribution to the binding affinity. Hot-site regions present particularly dense concentrations of contacts. We want to verify whether these hot-sites might be one or several, and where they are preferentially located. This problem is far from trivial and indeed no geometrical characterisation of interaction signals has been proposed in the description of binding affinity so far, to the best of our knowledge. One main reason is that the molecular structure does not easily identify non-covalent interactions and that a fine simultaneous exploration of a wide range of non-covalent interaction types is necessary (Keinan et al. 2004, Johnson et al. 2010).

In LISA, we exploit the Non-Covalent Interaction (NCI) approach (Johnson et al. 2010), based on electron density and its derivatives, to reveal the underlying chemistry that complements covalent structure. NCI uses only knowledge on the atomic coordinates to represent non-bonded interactions as continuous surfaces rather than close contacts between atom pairs. These surfaces are used to identify the favorable and non-favorable interactions for a complex and quantify their strength. Namely, low/high electron density values appear as indicator of weak/strong interaction strengths (Gibbs et al. 2004) and, in turn, stronger interactions are identified either as stabilising/favorable or de-stabilising/non-favorable interactions by density derivatives. See **Figure 1** (and STAR Methods).

To include the information on favorable and non-favorable interactions in our geometrical analysis of the protein interfaces, we represent the interface in 3 dimensions by an enveloping "cuboid grid" (called "LISA cuboid grid"; see STAR Methods and **Figure 2**). The enveloping cube is divided in 125 (5^3) smaller "LISA cubes" where the counting of the number of favorable contacts and their strength (corresponding to ϱ - see STAR Methods) is realised locally. The set of cubes is classified in three subsets depending on the strength of the contacts localised in them, where cubes with highest strength identify the hot-sites of the interaction.

Three more characteristics of the interface are explored in our evaluation of potential interactions. First, we analyse the interaction surfaces of both partners with the Support Core Rim (SCR) model proposed in (Levy 2010). SCR describes the experimental protein-protein interaction sites as comprising

three concentric layers: a central layer comprised of mostly buried residues (Support), an intermediate layer of surface residues that become buried upon association with the partner (Core), and an outer layer of surface residues remaining partially exposed to the solvent in the complex (Rim). In (Laine & Carbone 2015), we demonstrated that the three layers satisfy different geometrical, physico-chemical and evolutionary (conservation) properties, and can be detected with high accuracy towards binding sites prediction. Taking into consideration the SCR structure of the two binding partners, one can localise favorable and non-favorable contacts in specific SCR regions. LISA model should confirm whether some contact class localised in a specific region plays a special role in the prediction of binding affinity.

Second, it has been shown that residues lying on the surface of the complex, and not in the interaction sites, influence the binding affinity (Kastritis et al. 2014, Marillet et al. 2016). We investigate which properties of the surface residues are most relevant for the interaction and we explore whether their nature is polar, apolar or charged. For this, we use the measure of Non-Interacting Surface (NIS) introduced in (Kastritis et al. 2014).

Third, based on DSSP (Define Secondary Structure of Proteins (Kabsch & Sander 1983*a*)) analysis, we investigate whether the presence of specific secondary structure elements at the interface might influence the binding affinity, and we verify whether there are preferable secondary structure elements lying at the interaction site.

By coupling the analysis of the distribution of favorable and non-favorable contacts, realised with NCI (Johnson et al. 2010), in the LISA cuboid grid with the SCR model, the DSSP analysis of the secondary structures of the interaction surfaces (Kabsch & Sander 1983*a*), and the NIS analysis of polar/apolar/charged residues lying at the complex surface (Kastritis et al. 2014, Marillet et al. 2016), LISA model associates to the interaction sites a score that corresponds to a non-linear combination of 10 features (obtained with a feature selection algorithm starting from about 200 features) coming from the four analyses, NCI, SCR, NIS and DSSP. LISA score is computed with a function issued by Support Vector Regression (Drucker et al. 1997) (on a radial kernel). The 10 features are the most important ones characterising protein-protein interaction surfaces and the score estimates the binding affinity of the two proteins. LISA flow is represented in **Figure 2** and LISA model is explained below.

LISA radial model

Based on the four analyses of a complex, NCI, SCR, DSSP and NIS, we defined 179 descriptors of protein-protein interfaces, including information on the total number of contacts, their strength, their location with respect to the SCR model, their concentration in hot-sites of high strength, their localisation on specific secondary structure elements (see STAR Methods). We also added the contribution coming from polar/apolar/charged surface residues lying outside of the interface. A feature selection algorithm, applied to a subset of 98 complexes from AffinityB2_rel, which we call Kastritis_rel, extracted 16 key features from the 179 descriptors. We further selected 10 features out of the 16 ones by running a Support Vector Regression (based on a radial kernel). The Kastritis_rel dataset was used as a training set, and the remaining 27 complexes from AffinityB2_rel were used as testing set, called Vreven_rel. 65 535 regression models were trained and tested. They correspond to all possible combinations of $n = 1 \dots 16$ of the 16 features. The best feature combination was chosen by maximizing both the correlation with the training set (Kastritis_rel) and the testing set (Vreven_rel). (See STAR Methods and **Supplemental Figure S3, top**.) We note that the Vreven_rel dataset was used only to select the 10 features retained to define the final model, but it was not used to train the parameters of that model. Moreover, our optimization strategy for model selection, using both Kastristis_rel and Vreven_rel datasets, was designed to yield a

more generic/transferable model than if we were to optimize on Kastristis_rel only, and it allows us to limit overtraining: **Supplemental Figure S3, top**, illustrates the performance of LISA model compared to the 100 best ones (see STAR Methods).

The resulting non linear function, called *LISA score*, is defined on the 10 selected features. An internal 4-fold cross-validation procedure applied to the complexes of the Kastritis_rel dataset (see STAR Methods and **Supplemental Table S1**) allowed us to estimate a Root Mean Square Error (RMSE) of 2.35 ± 0.38 kcal/mol. (See STAR Methods.)

The 10 features are reported in **Supplemental Table S1**. Most of them describe a strong contribution played by core-core contacts and a minor contribution played by support (support-support and support-rim) contacts. Rim contacts appear to bring a marginal contribution. More precisely, the total number of core-core contacts (Tot$|CC$) and the total strength of core-core contacts localised in medium strength regions (Strength$|CC|IS_2$, where IS_2 denotes regions with medium strength, see STAR Methods) provide a first strong contribution to the interaction. The strength of unfavorable core-core contacts in weak strength regions (Strength$_{\text{unfav}}|CC|IS_3$, where IS_3 denotes regions with weak strength, see STAR Methods) and, to a minor extent, the concentration of favorable core-core contacts in high strength regions (Strength$_{\text{fav}}|CC|IS_1$, where IS_1 denotes regions with high strength, see STAR Methods) complete the influence of contacts contribution in core regions. In the support region of the interaction site, two other features highlight the positive contribution of support-support contacts (Strength$|SS|IS_2$) and favorable support-rim contacts (Strength$_{\text{fav}}|RS|IS_2$) localised in medium strength regions. One more feature highlights the presence of 3_{10}-helix structural elements at the interface. Among the features that are anti-correlated with binding affinity, identifying characteristics that tend to be avoided by high affinity complexes, we find the α-helix structural element, a high strength of RR favorable contacts (Strength$_{\text{fav}}|RR$), and, especially, a high percentage of charged residues lying at the complex surface (NIS_{charged}) negatively contribute to binding affinity.

All our analyses, NCI, SCR, NIS and DSSP, contribute to the 10 features.

An important outcome of this geometrical analysis of protein interfaces concerns the distribution of interaction signals in interaction surfaces. A schema suggesting the tendency for (favorable) contacts to be distributed on specific regions of the protein interaction surface is given in **Figure 3**.

Correlation between LISA and experimental binding affinity

Different experimental methods can be used to determine binding free energy and the results should be interpreted by taking into account the sensitivity and the strength of the employed experimental technique. (See **Data S1** for the detailed list of techniques.) In fact, as already reported in (Vangone & Bonvin 2015), the choice of the experimental technique can greatly influence the correlation with the predictions. LISA records R = 0.74 ($p < 1.023e$-06) with ITC (32 complexes), R = 0.84 ($p < 2.2e$-16) with SPR (60 complexes), R = 0.77 ($p = 0.005856$) with stopped-flow fluorimetry (11 complexes), and R = 0.75 ($p = 6.1e$-05) with spectroscopy (22 complexes) measures. It records a correlation R=0.81 ($p < 2.2e$-16) with the full set AffinityB2_rel experimental measures (**Figure 4, top left**, and **Supplemental Table S2, top**). Note that the set of complexes measured with techniques considered as non-reliable such as inhibition_assay (see STAR Methods) shows R = 0.20 ($p < 0.15$, **Figure 4, bottom left**). The analysis of all complexes in AffinityB2, experimentally evaluated with reliable and non-reliable techniques, gives $R = 0.66$.

When the analysis is realised on the Katristis_rel dataset, made of 98 complexes and used to train LISA, the correlation is $R = 0.84$ ($p < 2.2e$-16; **Figure 4, top right**). A slight improvement of $R = 0.86$

($p < 2.2e$-16) is obtained on the Katritis81 database used in (Vangone & Bonvin 2015) and made of 81 complexes experimentally studied with reliable experimental techniques. Note that on Katritis81, LISA records R = 0.75 ($p < 1.357e$-04) with ITC (20 complexes), R = 0.86 ($p < 2.753e$-12) with SPR (39 complexes), R = 0.86 ($p = 0.005794$) with stopped-flow fluorimetry (8 complexes), and R = 0.86 ($p = 6.796e$-05) with spectroscopy (14 complexes). See **Figure 4, bottom right**. For each technique, note that the correlations obtained by Prodigy (Vangone & Bonvin 2015) on the same complexes are comparable or lower: R = 0.78 for ITC, R = 0.69 for SPR, R = 0.68 for stopped-flow fluorimetry, and R = 0.75 for spectroscopic methods.

Comparison between LISA and 17 other predictors. The performance of LISA was compared with 17 state-of-the-art predictors, based on physical potentials and composite scoring functions. The performance of the four most successful tools (ZAPP, Prodigy, Firedock-EI, INSIDE) on AffinityB2_rel is reported in **Figure 5, top**, and, for all predictors, in **Supplemental Figure S1** (see also **Supplemental Table S2, top**). LISA reaches best correlations on experimental binding affinity with $R = 0.84$ and 0.67 for Kastritis_rel and Vreven_rel datasets, while all other approaches reach a correlation of at most 0.66 and 0.57 respectively, and, in general, much lower values. We note that the Vreven_rel dataset was not used for the training of any of the 17 predictors and that this unbiased evaluation shows that LISA's performance is robust on different subsets.

We observe that LISA outperforms all methods behaving consistently well on AffinityB2_rel ($R = 0.81$) and its subsets (**Figure 5, top**). By measuring the amplitude of the conformational changes that take place upon binding with Interface rmsd (I-rmsd; see STAR Methods), we labelled the AffinityB2_rel complex interfaces as rigid (I-rmsd < 1Å), medium-rigid ($1\mathring{A} \leq I - rmsd \leq 2$Å) and flexible (I-rmsd > 2Å) as in (Vreven et al. 2012) (see STAR Methods) and observed that LISA displays the same behaviour on rigid and medium-rigid complexes ($R = 0.81$ and 0.8 respectively), and a highly improved behaviour on flexible complexes with a correlation $R = 0.9$ compared to ZAPP, Prodigy, Firedock-EI and INSIDE scoring correlations ≤ 0.6 (**Figure 5, bottom**; see **Supplemental Table S3**). We also compared LISA performance on Kastritis81 and on AffinityB2_rel* to Prodigy and ConsBind, respectively. On these subsets, LISA performance remains consistent with the behaviour observed on AffinityB2_rel (**Figure 5, top**; see also **Supplemental Tables S4**, **S2 bottom** and **S3 bottom**). We note that on flexible and rigid complexes in Kastritis81, LISA reaches the very high correlation of $R = 0.87$.

When evaluated on AffinityB2_rel complex interfaces labelled as rigid (I-rmsd $\leq$ 1Å) and flexible (I-rmsd > 1Å) as in (Vangone & Bonvin 2015), LISA behaviour is globally stable, with $R = 0.8$ and 0.82 on the two sets, respectively. In contrast, the majority of the other 17 predictors display a sharper difference between rigid and flexible complexes than LISA, with a much lower performance on flexible ones (**Supplemental Figure S1** and **Supplemental Tables S5** and **S4**). This is probably due to the genericity of the geometric model LISA encodes, which highlights characteristics that may be "universal" for protein-protein interactions, as argued in the Discussion.

We also evaluated LISA on sets of complexes satisfying different functions. **Supplemental Table S6, top** (see also **Supplemental Table S6, bottom**) reports the correlations between binding energies obtained with different tools and the experimental binding energy associated to different functional classes. It is interesting to observe that there is no tool that outperforms the others in all classes. LISA performs very well on membrane receptors (OR, R=0.91), enzyme/regulatory subunits (ER, R=0.99), enzyme/substrates (ES, R=0.78) and G-protein containing (OG, R=0.88) interactions. It obtains the best performance on antibody-antigens (A, R=0.51) and others (OX, R=0.75) compared to all other predictors. It does not perform well on the set of 22 enzyme-inhibitors (EI, R=0.55) nor on the 4 antigen-bound antibodies (AB,

R=0.63). Overall, LISA outperforms 6 tools on all 8 functional classes, 10 tools on 7 classes and 1 tool on 6 classes.

Vreven_rel is disjoint from Kastritis_rel and provides an important reference test where to verify accuracy. In reality, this dataset of 27 complexes has a biased composition towards antibody-antigenes, a difficult functional class where LISA and all other computational approaches reach the worse performance (see **Supplemental Figure S1** and **Supplemental Table S2, top**). Moreover, it has no complex from AB, ES and OG classes, where LISA performs the best. Nevertheless, on this dataset, LISA performs successfully compared to the other approaches, with an overall correlation $R = 0.67$. Note that on this dataset, ZAPP drops to $R = 0.36$ and Prodigy to $R = 0.13$.

Evaluation of the four LISA modules. LISA features, based on information coming from the four different modules of the analysis, NCI, SCR, DSSP and NIS (**Figure 2**), have been analysed by discarding each module separately and by evaluating the performance of the resulting "simplified" method. Correlations of the predictions performed by each "simplified" method with experimental binding energies of complexes belonging to AffinityB2_rel, Kastritis_rel and Vreven_rel datasets are reported in **Table 1**. From the outcomes, the four modules NCI, SCR, DSSP and NIS play an important role in LISA, with the best performance reached with the contribution of all modules combined. The NCI module appears to be the most relevant for the three datasets, followed by SCR. This is especially visible in Vreven_rel analysis, where DSSP seems to play the smaller role. As observed before, the functional composition of the Vreven_rel dataset is biased, and our results suggest that the role of the modules might be dependent on specific functional classes.

LISA computational time. LISA is scalable and its bottleneck is NCI calculation. NCI analysis can take advantage of a multithreading computation though, as shown in **Supplemental Table S7**, and high performance computing (HPC) can drastically improve computational performance. Note that LISA was not designed for a large scale treatment of millions of decoys even though HPC can be used for a reasonably fast treatment of many complex conformations.

Discussion

LISA demonstrates that improvements in binding affinity prediction can be reached by modelling the geometry of the distribution of contacts at interacting surfaces. Besides considering secondary structures at the interface (DSSP) and the composition of the non-interacting surface (NIS), used in the past to estimate binding affinities, it introduces two features, namely the favorable/unfavorable quality of contacts, their strength and concentration in specific regions, both inferred from quantum chemistry calculations (NCI), and the support-core-rim model for protein-protein interfaces (SCR). With only 10 easily interpretable descriptors, LISA final model enables to learn about protein-protein association. It identifies the properties contributing the most to the binding (see **Figure 3** and the associated description) by showing the importance of the geometrical distribution of contacts and hot-sites at the interacting surface. These findings were not previously reported. LISA compares favourably to many other state-of-the-art scoring functions. Its computational bottleneck is NCI which performs the quantum chemistry calculations, but the computational efficiency of this step could be improved and work on this direction is undergoing.

3D localization of non-covalent interactions. Protein complexes do not easily identify the intricate non-covalent interactions governing their binding. These interactions occupy broader regions in space than

close contacts between atom pairs. These regions can be represented by continuous surfaces (Johnson et al. 2010), where the density of the signal and the strength of the interaction can be numerically evaluated based on the electron density and its derivatives. We demonstrate that this fine description of the interaction, taking into consideration all small non-covalent contributions and their spatial distribution, goes far beyond the usual residue-residue contacts considered in other computational approaches and leads to define a more realistic model of binding affinity improving our understanding of protein-protein interactions.

From atom-atom contacts to isosurfaces and back to contacts again. We wish to stress that, in LISA, the properties assigned to a contact and used to evaluate binding affinity are not intrinsic to the contact itself but coming from a global analysis of the interaction between the contact and its environment, i.e. the other contacts around it. This is made possible by NCI analysis which generates isosurfaces representing the effects of multiple atomic interactions. Each isosurface is a discrete object, namely a region of the NCI grid. This enables to associate to each interatomic contact the local values of the isosurface it intersects, corresponding to the closest point in the NCI grid.

Contribution of favorable contacts in binding affinity. LISA demonstrates that by coupling favorable non-covalent interactions with the geometry of the interface (described by the Support-Core-Rim model (Levy 2010)) and the residues on the complex surface influencing the interaction turns out to be crucial to improve previous models of binding affinity.

From the LISA model, we learn that the number and strength of CC contacts are crucial for predicting binding affinity. In particular, we learn that favorable contacts are more important than generic contacts and that, in general, some regions where the contacts are highly concentrated (hot-sites), depending on their localisation at the surface, are more important than others. This is the case for favorable CC, RR and RS contacts. Noticeably, unfavorable CC contacts concentrated in regions of very weak strength also play a role in binding. The spread distribution of contacts in the core and of favorable contacts in the rim regions of the interfaces leads to high binding affinity. A minor role is played by secondary structure elements with 3_{10}-helix motifs favouring the interaction.

LISA model was defined and tested on a broad set of complexes, for which we expect binding behaviour to be rather different ensuring genericity of the model. Nevertheless, we can expect that some variations on the model might take place when the number of experimental complexes will be augmented. **Supplemental Figure S3** highlights that the 100 best scoring models share a number of features. Some other features are shared by a large proportion of the models and we expect that, with new experimental measures, some of those features might become preferential. The features, being present in only a few models may be important to characterize the binding of complexes involved in specific functions.

LISA on flexible complexes. LISA displays a comparable performance on rigid and flexible complexes, while all other predictors see their performance drop on flexible complexes. This result raises some questions on the general character of the descriptors used in LISA to predict binding affinity. Indeed, it suggests that the features used by all other tools are less significant on flexible complexes than on rigid ones, while, in contrast, LISA properties appear to be more "universal".

The adaptability of LISA interface descriptors to complexes formed by proteins displaying different conformational adjustments and possibly large conformational changes upon binding together, suggest that these features might be driving the conformational change and it could be of interest to use them for studying the dynamics of complex formation.

Improving LISA model. One might wonder how we can improve further LISA model. Certainly, solvatation is a missing component of the LISA model, and one should consider the role of solvent as a mediator of the interactions at the interface. Since water molecules can realise bridges between residues, an explicit inclusion of them in the model could improve it. More generally, one can expect that more reliable experimental measures on new complexes will enable to improve model fitting and validation.

Can binding affinity play a role in discriminating partners? With LISA, we analysed the distribution of geometrical features of atom-atom contacts in a complex interface and demonstrated that specific characteristics of this distribution become helpful to estimate the binding energy. In this respect, it is important to notice that the geometrical distribution of interaction signals seems to provide more insights on true partners than the binding energy itself. Indeed, molecular docking highlighted that the binding energy alone is not sufficient to discriminate true partners from false ones but that knowing the precise location of the binding site helps to discriminate partners (Sacquin-Mora et al. 2008, Lopes et al. 2013). Also, docking procedures highlighted that there is a tendency for partners and non-partners to preferably dock the same sites (Fernandez-Recio et al. 2004, Sacquin-Mora et al. 2008, Martin & Lavery 2012, Lopes et al. 2013, Laine & Carbone 2016, Vamparys et al. 2016). These two observations lead to draw a special attention on the way atom-atom contacts are established between real partners compared to non-real ones. Indeed, given the same interaction site, one expects contact distributions to be different for distinguished partners and among different conformations of the same partners. In the future, we wish to show that LISA model, based on a geometrical distribution of specific favorable contacts at the interface, could be included in interaction indices to better discriminate partners from non-partners.

A direction of investigation for design. In (Johnson et al. 2010), surfaces describing non-covalent interactions are proposed as offering rich insight into the design of improved ligands. Here, in a similar manner, we propose to take into account all the small contributions justifying the ligand fit on the specific binding site for evaluating their binding affinity. Further investigations of the geometry of the isosurfaces will likely offer insights on our understanding of molecular interactions.

Acknowledgements

We thank Anna Vangone for useful information on the way to use Prodigy and for providing the tool, Julia Contreras-Garcia for help in running NCI, Juliana Bernardes for suggestions on SVR and Chloé Dequeker for a script to visualize LISA contacts. We acknowledge the access to the HPC resources of the Institute for Scientific Computing and Data at UPMC (Equip@Meso project - ANR-10-EQPX- 29-01, Excellence Program "Investissement d'Avenir"). Funding: Excellence Programme "Investissement d'Avenir" en Bioinformatique - Ministère de l'Enseignement Supérieur et de la Recherche, France (MAPPING - ANR-11-BINF-0003 to AC); Institut Universitaire de France (AC); LabEx CALSIMLAB - Programme "Investissement d'Avenir" - Ministère de l'Enseignement Supérieur et de la Recherche, France (CALSIMLAB - ANR-11-LABX-0037-01 to RR, AC). The funder had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

RR: implementation of LISA software; EL, RR: performance of computational analysis; AC, RR: conception and design, analysis and interpretation of data; AC: writing of the article.

Declaration of interests

The authors declare no competing interests.

References

Aloy, P. & Russell, R. B. (2006), 'Structural systems biology: modelling protein interactions', *Nature Reviews Molecular Cell Biology* **7**(3), 188–197.

Andrusier, N., Nussinov, R. & Wolfson, H. J. (2007), 'FireDock: fast interaction refinement in molecular docking', *Proteins: Structure, Function, and Bioinformatics* **69**(1), 139–159.

Audie, J. & Scarlata, S. (2007), 'A novel empirical free energy function that explains and predicts protein–protein binding affinities', *Biophysical Chemistry* **129**(2), 198–211.

Bai, H., Yang, K., Yu, D., Zhang, C., Chen, F. & Lai, L. (2011), 'Predicting kinetic constants of protein–protein interactions based on structural properties', *Proteins: Structure, Function, and Bioinformatics* **79**(3), 720–734.

Becke, A. D. (1995), 'Exchange-correlation approximations in density-functional theory.', *Advanced Series in Physical Chemistry 2 (Modern Electronic Structure Theory, Pt. 2)* .

Beltrao, P., Kiel, C. & Serrano, L. (2007), 'Structures in systems biology', *Current Opinion in Structural Biology* **17**(3), 378–384.

Ben-Shimon, A. & Eisenstein, M. (2010), 'Computational mapping of anchoring spots on protein surfaces', *Journal of Molecular Biology* **402**(1), 259–277.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S. & et al. (2002), 'The protein data bank', *Acta Crystallographica Section D Biological Crystallography* **58**(6), 899–907.

Betzi, S., Restouin, A., Opi, S., Arold, S. T., Parrot, I., Guerlesquin, F., Morelli, X. & Collette, Y. (2007), 'Protein–protein interaction inhibition (2P2I) combining high throughput and virtual screening: application to the HIV-1 Nef protein', *Proceedings of the National Academy of Sciences* **104**(49), 19256–19261.

Chaudhury, S., Lyskov, S. & Gray, J. J. (2010), 'PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta', *Bioinformatics* **26**(5), 689–691.

Chen, J., Sawyer, N. & Regan, L. (2013), 'Protein-protein interactions: General trends in the relationship between binding affinity and interfacial buried surface area', *Protein Science* **22**(4), 510–515.

Cohen, A. J., Mori-Sánchez, P. & Yang, W. (2008), 'Insights into current limitations of density functional theory', *Science* **321**(5890), 792–794.

Contreras-Garcia, J., Johnson, E. R., Keinan, S., Chaudret, R., Piquemal, J.-P., Beratan, D. N. & Yang, W. (2011), 'Nciplot: A program for plotting noncovalent interaction regions', *Journal of Chemical Theory and Computation* **7**(3), 625–632.

Dell'Orco, D. (2009), 'Fast predictions of thermodynamics and kinetics of protein–protein recognition from structures: from molecular design to systems biology', *Molecular Biosystems* **5**(4), 323–334.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., Vapnik, V. et al. (1997), 'Support vector regression machines', *Advances in Neural Information Processing Systems* **9**, 155–161.

Erijman, A., Rosenthal, E. & Shifman, J. M. (2014), 'How structure defines affinity in protein-protein interactions', *PLOS One* **9**(10), e110085.

Fernandez-Recio, J., Totrov, M. & Abagyan, R. (2004), 'Identification of protein-protein interaction sites from docking energy landscapes', *J. Mol. Biol.* **335**(3), 843–865.

Fleishman, S. J., Whitehead, T. A., Strauch, E.-M., Corn, J. E., Qin, S., Zhou, H.-X., Mitchell, J. C., Demerdash, O. N., Takeda-Shitaka, M., Terashi, G. et al. (2011), 'Community-wide assessment of protein-interface modeling suggests improvements to design methodology', *Journal of Molecular Biology* **414**(2), 289–302.

Gibbs, G. V., Cox, D. F. & Rosso, K. M. (2004), 'A connection between empirical bond strength and the localization of the electron density at the bond critical points of the sio bonds in silicates', *The Journal of Physical Chemistry A* **108**(38), 7643–7645.

Grosdidier, S., Pons, C., Solernou, A. & Fernández-Recio, J. (2007), 'Prediction and scoring of docking poses with pydock', *Proteins: Structure, Function, and Bioinformatics* **69**(4), 852–858.

Gunther, D., Boto, R. A., Contreras-Garcia, J., Piquemal, J.-P. & Tierny, J. (2014), 'Characterizing molecular interactions in chemical systems', *IEEE Transactions on Visualization and Computer Graphics* **20**(12), 2476–2485.

Hakes, L., Pinney, J. W., Robertson, D. L. & Lovell, S. C. (2008), 'Protein-protein interaction networks and biology - what's the connection?', *Nature Biotechnology* **26**(1), 69–72.

Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. (1999), 'From molecular to modular cell biology', *Nature* **402**, C47–C52.

Hohenberg, P. & Kohn, W. (1964), 'Inhomogeneous electron gas', *Physical review* **136**(3B), B864.

Horton, N. & Lewis, M. (1992), 'Calculation of the free energy of association for protein complexes', *Protein Science* **1**(1), 169–181.

Hubbard, S. J. & Thornton, J. M. (1993), 'Naccess', *Computer Program, Department of Biochemistry and Molecular Biology, University College London* **2**(1).

Jiang, L., Gao, Y., Mao, F., Liu, Z. & Lai, L. (2002), 'Potential of mean force for protein–protein interaction studies', *Proteins: Structure, Function, and Bioinformatics* **46**(2), 190–196.

Jiménez-García, B., Pons, C. & Fernández-Recio, J. (2013), 'pyDoakWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring', *Bioinformatics* p. btt262.

Johnson, E. R., Keinan, S., Mori-Sanchez, P., Contreras-Garcia, J., Cohen, A. J. & Yang, W. (2010), 'Revealing noncovalent interactions', *Journal of the American Chemical Society* **132**(18), 6498–6506.

Jones, S. & Thornton, J. M. (1996), 'Principles of protein-protein interactions', *Proceedings of the National Academy of Sciences* **93**(1), 13–20.

Joosten, R. P., Te Beek, T. A., Krieger, E., Hekkelman, M. L., Hooft, R. W., Schneider, R., Sander, C. & Vriend, G. (2011), 'A series of pdb related databases for everyday needs', *Nucleic Acids Research* **39**(suppl 1), D411–D419.

Kabsch, W. & Sander, C. (1983*a*), 'Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features', *Biopolymers* **22**(12), 2577–2637.

Kabsch, W. & Sander, C. (1983*b*), 'Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features', *Biopolymers* **22**(12), 2577–2637.

Kastritis, P. L. & Bonvin, A. M. (2010), 'Are scoring functions in protein-protein docking ready to predict interactomes? Clues from a novel binding affinity benchmark', *Journal of Proteome Research* **9**(5), 2216–2225.

Kastritis, P. L., Moal, I. H., Hwang, H., Weng, Z., Bates, P. A., Bonvin, A. M. J. J. & Janin, J. (2011), 'A structure-based benchmark for protein-protein binding affinity', *Protein Science* **20**(3), 482–491.

Kastritis, P. L., Rodrigues, J. P., Folkers, G. E., Boelens, R. & Bonvin, A. M. (2014), 'Proteins feel more than they see: Fine-tuning of binding affinity by properties of the non-interacting surface', *Journal of Molecular Biology* **426**(14), 2632–2652.

Keinan, S., Ratner, M. A. & Marks, T. J. (2004), 'Molecular zippers–designing a supramolecular system', *Chemical Physics Letters* **392**(4), 291–296.

Keskin, O., Ma, B., Rogale, K., Gunasekaran, K. & Nussinov, R. (2005), 'Protein–protein interactions: organization, cooperativity and mapping in a bottom-up systems biology approach', *Physical Biology* **2**(2), S24.

Kiel, C., Beltrao, P. & Serrano, L. (2008), 'Analyzing protein interaction networks using structural information', *Annu Rev Biochem* **77**, 415–441.

Kollman, P. A. (1977), 'Noncovalent interactions', *Acc. Chem. Res* **10**(10), 365–371.

Kortemme, T., Joachimiak, L. A., Bullock, A. N., Schuler, A. D., Stoddard, B. L. & Baker, D. (2004), 'Computational redesign of protein-protein interaction specificity', *Nature Structural & Molecular Biology* **11**(4), 371–379.

Laine, E. & Carbone, A. (2015), 'Local geometry and evolutionary conservation of protein surfaces reveal the multiple recognition patches in protein-protein interactions', *PLoS Comput Biol* **11**(12), e1004580.

Laine, E. & Carbone, A. (2016), 'Protein social behavior makes a stronger signal for partner identification than surface geometry', *Proteins: Structure, Function, and Bioinformatics* .

Levy, E. D. (2010), 'A simple definition of structural regions in proteins and its use in analyzing interface evolution', *Journal of Molecular Biology* **403**(4), 660–670.

Liu, S., Zhang, C., Zhou, H. & Zhou, Y. (2004), 'A physical reference state unifies the structure-derived potential of mean force for protein folding and binding', *Proteins: Structure, Function, and Bioinformatics* **56**(1), 93–101.

Lopes, A., Sacquin-Mora, S., Dimitrova, V., Laine, E., Ponty, Y. & Carbone, A. (2013), 'Protein-protein interactions in a crowded environment: an analysis via cross-docking simulations and evolutionary information', *PLoS Comput Biol* **9**(12), e1003369.

Ma, X. H., Wang, C. X., Li, C. H. & Zu Chen, W. (2002), 'A fast empirical approach to binding free energy calculations based on protein interface information', *Protein Engineering* **15**(8), 677–681.

Marillet, S., Boudinot, P. & Cazals, F. (2016), 'High-resolution crystal structures leverage protein binding affinity predictions', *Proteins: structure, function, and bioinformatics* **84**(1), 9–20.

Martin, J. & Lavery, R. (2012), 'Arbitrary protein-protein docking targets biologically relevant interfaces', *BMC Biophys* **5**, 7.

McGuffee, S. R. & Elcock, A. H. (2010), 'Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm', *PLoS Comput Biol* **6**(3), e1000694.

Moal, I. H., Agius, R. & Bates, P. A. (2011), 'Protein–protein binding affinity prediction on a diverse set of structures', *Bioinformatics* **27**(21), 3002–3009.

Moal, I. H., Jiménez-García, B. & Fernández-Recio, J. (2015), 'CCharPPI web server: computational characterization of protein–protein interactions from structure', *Bioinformatics* **31**(1), 123–125.

Nooren, I. M. & Thornton, J. M. (2003), 'Structural characterisation and functional significance of transient protein–protein interactions', *Journal of molecular biology* **325**(5), 991–1018.

Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. (2010), 'Transient protein-protein interactions: structural, functional, and network properties', *Structure* **18**(10), 1233–1243.

Pierce, B. & Weng, Z. (2007), 'ZRANK: reranking protein docking predictions with an optimized energy function', *Proteins: Structure, Function, and Bioinformatics* **67**(4), 1078–1086.

Pierce, B. & Weng, Z. (2008), 'A combination of rescoring and refinement significantly improves protein docking performance', *Proteins: Structure, Function, and Bioinformatics* **72**(1), 270–279.

Pons, C., Talavera, D., de la Cruz, X., Orozco, M. & Fernandez-Recio, J. (2011), 'Scoring by intermolecular pairwise propensities of exposed residues (SIPPER): A new efficient potential for protein- protein docking', *Journal of Chemical Information and Modeling* **51**(2), 370–377.

R Development Core Team (2011), *R: A Language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Ravikant, D. & Elber, R. (2010), 'Pie–efficient filters and coarse grained potentials for unbound protein–protein docking', *Proteins: Structure, Function, and Bioinformatics* **78**(2), 400–419.

Sacquin-Mora, S., Carbone, A. & Lavery, R. (2008), 'Identification of protein interaction partners and protein-protein interaction sites', *J. Mol. Biol.* **382**(5), 1276–1289.

Sharabi, O., Yanover, C., Dekel, A. & Shifman, J. M. (2011), 'Optimizing energy functions for protein–protein interface design', *Journal of Computational Chemistry* **32**(1), 23–32.

Su, Y., Zhou, A., Xia, X., Li, W. & Sun, Z. (2009), 'Quantitative prediction of protein–protein binding affinity with a potential of mean force considering volume correction', *Protein Science* **18**(12), 2550–2558.

Tian, F., Lv, Y. & Yang, L. (2012), 'Structure-based prediction of protein–protein binding affinity with consideration of allosteric effect', *Amino Acids* **43**(2), 531–543.

Ubbink, M. (2009), 'The courtship of proteins: understanding the encounter complex', *Febs Letters* **583**(7), 1060–1066.

Vamparys, L., Laurent, B., Carbone, A. & Sacquin-Mora, S. (2016), 'Great interactions: How binding incorrect partners can teach us about protein recognition and function', *Proteins: Structure, Function, and Bioinformatics* **84**(10), 1408–1421.

Vangone, A. & Bonvin, A. M. (2015), 'Contacts-based prediction of binding affinity in protein-protein complexes', *eLife* **4**.

Viswanath, S., Ravikant, D. & Elber, R. (2013), 'Improving ranking of models for protein complexes with side chain modeling and atomic potentials', *Proteins: Structure, Function, and Bioinformatics* **81**(4), 592–606.

Vreven, T., Hwang, H., Pierce, B. G. & Weng, Z. (2012), 'Prediction of protein-protein binding free energies', *Protein Science* **21**(3), 396–404.

Vreven, T., Moal, I. H., Vangone, A., Pierce, B. G., Kastritis, P. L., Torchala, M., Chaleil, R., Jimenez-Garcia, B., Bates, P. A., Fernandez-Recio, J. & et al. (2015), 'Updates to the integrated protein–protein interaction benchmarks: Docking Benchmark version 5 and Affinity Benchmark version 2', *Journal of Molecular Biology* **427**(19), 3031–3041.

Yang, Y. & Zhou, Y. (2008*a*), 'Ab initio folding of terminal segments with secondary structures reveals the fine difference between two closely related all-atom statistical energy functions', *Protein Science* **17**(7), 1212–1219.

Yang, Y. & Zhou, Y. (2008*b*), 'Specific interactions for ab initio folding of protein terminal regions with secondary structures', *Proteins: Structure, Function, and Bioinformatics* **72**(2), 793–803.

Zeng, Z.-H. & Li, Y. C. (2008), 'Empirical parameters for estimating protein-protein binding energies: number of short-and long-distance atom-atom contacts', *Protein and Peptide Letters* **15**(2), 223–231.

Zhao, L. & Chmielewski, J. (2005), 'Inhibiting protein–protein interactions using designed molecules', *Current Opinion in Structural Biology* **15**(1), 31–34.

Zhou, P., Wang, C., Tian, F., Ren, Y., Yang, C. & Huang, J. (2013), 'Biomacromolecular quantitative structure–activity relationship (bioqsar): a proof-of-concept study on the modelling, prediction and interpretation of protein–protein binding affinity', *Journal of computer-aided molecular design* **27**(1), 67–78.

Main figures titles and legends

Figure 1. Analysis of atom-atom contacts in non-covalent interactions. Images are realised with NCIPLOT (Johnson et al. 2010). **A.** NCI grid representing the contact surface between two proteins (pdb 1E6E). (The NCI grid is defined with no filter on RDG values.) **B.** Isosurfaces characterising regions with lowest density gradient, $s \leq 1$, within the contact surface in **A**. **C.** Analysis of favorable vs non-favorable contact regions. Gradient isosurfaces in **B** are coloured on a blue-green scale according to the value of $\text{sign}(\lambda_2) \times \rho$: blue indicates favorable interactions ($\varrho \leq 0$) and green indicates unfavourable ones ($\varrho > 0$). Atom-atom contacts intersecting favorable and non-favorable contact regions are colored red.

Figure 2. LISA flow. LISA flow is constructed around four blocks of analysis of the protein complex detecting different signals: NCI (identification of favorable and non-favorable contacts and their strength), SCR (positioning of the atom-atom contacts within the geometrical SCR model), DSSP (identification of the secondary structures at the interaction surface) and NIS (evaluation of the influence of polar/ apolar/ charged residues lying on the complex surface but not belonging to the interface). Signals coming from NCI, SCR, DSSP and NIS analysis are combined together in a non-linear combination defining a global score for the interface. LISA geometrically organises all these signals together in a grid (the LISA cuboid grid) highlighting the hot-sites of the interaction within a complex. The role of each module included in LISA and making a contribution to the description of the binding affinity is shown in Table 1.

Figure 3. Schema of contact distribution in protein interacting surfaces. Left: The interacting surface is structured in the three SCR layers: the support (S) is localised in the center (white), the core (C) is the central concentric layer (light brown) and the rim (R) is the most external layer (dark grey). Contacts (crosses) and favorable contacts (circles) are shown to distribute mostly on the core region in high (blue) and medium (green) concentration hot-sites. Medium concentration hot-sites contribute to the binding from the three (support, core and rim) regions. Unfavorable contacts in the core region play also a role. **Right:** atom-atom contacts between two interacting proteins (pdb 3SGB). An interacting protein surface is represented by sticks (only interacting residues are shown) and the other by balls. Residues (sticks and balls) are colored in white, light brown and dark grey depending on whether they belong to support, core or rim. Atom-atom contacts have been colored based on LISA analysis (see `contacts_details.txt` file generated by LISA): favorable contacts ($\rho < 0$) in IS1 (`IS1=1`) are blue, favorable contacts in IS2 (`IS2=1`) are green and unfavorable contacts in IS3 (`IS1=IS2=2`) are black.

Figure 4. Correlation of LISA scores with binding affinities, organised by experimental techniques. Top left: correlation between LISA scores on the 125 complexes in AffinityB2_rel and their equilibrium dissociation constant (Vreven et al. 2015) obtained by reliable experimental techniques: ITC, SPR, spectroscopy and stopped-flow. The associated p-value is also reported: R=0.81 and $p < 2.2e\text{-}16$. LISA is trained on Kastritis_rel, a subset of AffinityB2_rel. See also **Supplemental Figure S2** for an organisation of the complexes on functional classes. **Bottom left**: correlation computed on complexes whose experimental measures were obtained with non-reliable experimental techniques: inhibition assays and others; R=0.20 and $p < 0.15$. LISA is trained on Kastritis_rel. **Top right**: correlation computed on complexes whose experimental binding affinity is measured on the Kastritis_rel dataset; R=0.84 and $p < 2.2\text{-}16$. **Bottom right**: correlation computed on complexes in Kastritis81, the subset of AffinityB2_rel

studied in (Vangone & Bonvin 2015); R=0.86 and $p < 2.2e$-16.

Figure 5. LISA performance and comparison with 5 other tools on AffinityB2_rel datasets and Kastritis81. Top: correlations between the binding affinity predicted by LISA and other tools with the experimental binding energies reported in (Vreven et al. 2015). LISA, ZAPP (Vreven et al. 2012), Prodigy (Vangone & Bonvin 2015), FireDock_EI (Andrusier et al. 2007) and INSIDE (Andrusier et al. 2007) have been evaluated on AffinityB2_rel, Kastritis_rel and Vreven_rel datasets (left; **Supplemental Table S2, top**). LISA was compared to Prodigy also on Kastritis81 (Vangone & Bonvin 2015) (center; **Supplemental Table S4**). The comparison with ConsBind was realised on Vreven_rel and the subsets of AffinityB2_rel, Katritis_rel, called AffinityB2_rel* and Kastritis_rel*, defined by eliminating complexes 1DE4, 1IQD, 1M10, 1NB5, 1NCA, 1NSN, 1UUG from AffinityB2_rel and Kastritis_rel datasets since they are missed in ConsBind evaluation. **Bottom**: correlations of scores computed by LISA and other tools with experimental binding affinities reported in (Vreven et al. 2015). Results are organised according to flexibility of the complexes for AffinityB2_rel (left; **Supplemental Table S3, top**), Kastritis81 (center; **Supplemental Table S4**) and AffinityB2_rel* (right; **Supplemental Table S3, bottom**). Kastritis81 and AffinityB2_rel* are described as for the top figure. The subsets rigid (I-rmsd $<$ 1Å), medium-rigid (1Å$\leq$ I-rmsd $\leq$ 2Å) and flexible (I-rmsd $>$ 2Å) are labeled '*' for Kastritis81 and '**' for AffinityB2_rel*. For LISA comparison with 17 tools, see **Supplemental Figure S1**.

Figure 6. Functional and structural compositions of AffinityB2_rel, Kastritis_rel and Vreven_rel datasets. A. Functional composition (in percentages) of the AffinityB2_rel dataset made of 125 protein complexes (Vreven et al. 2015): antibodies (A), bound antibodies (AB), enzyme/inhibitors (EI), enzyme/substrate (ES), enzyme/regulatory subunit (ER), G-protein containing (OG), membrane receptors (OR), and miscellaneous (OX). **B.** Classification of AffinityB2_rel complexes in rigid (I-rmsd $<$ 1Å), medium-rigid (1Å$\leq$ I-rmsd $\leq$ 2Å) and flexible (I-rmsd $>$ 2Å) structures. The exact proportions and exact number of complexes for the subsets in A and B are reported in **Supplemental Table S8**.

Main tables and legends

	LISA/NCI	LISA/SCR	LISA/DSSP	LISA/NIS	LISA
NCI	-	+	+	+	+
SCR	+	-	+	+	+
DSSP	+	+	-	+	+
NIS	+	+	+	-	+
AffinityB2_rel	0.65	0.69	0.75	0.73	**0.81**
Kastritis_rel	0.74	0.75	0.78	0.77	**0.84**
Vreven_rel	0.28	0.40	0.63	0.48	**0.67**
RMSE	2.20 ± 0.42	2.24 ± 0.41	2.27 ± 0.41	2.35 ± 0.38	2.35 ± 0.38

Table 1. Contribution of the four main block analyses in LISA. The four blocks of analysis constituting the LISA model (NCI, SCR model, DSSP and NIS) are analysed by dropping them, one by one, from the LISA model to evaluate their contribution in LISA. The Root Mean Square Error (RMSE) is reported for each analysis of the Kastritis_rel dataset.

STAR Methods

Datasets. LISA performance was tested on three datasets. We considered the Affinity Benchmark version 2 (called 'AffinityB2' for short) (Vreven et al. 2015) comprising 144 complexes from (Kastritis et al. 2011) (called 'Kastritis') and 35 new complexes from (Vreven et al. 2015) (called 'Vreven'). The binding free energy reported in (Vreven et al. 2015) for the 179 complexes in AffinityB2 was estimated by specific technologies: stopped-flow fluorimetry, surface plasmon resonance/SPR, isothermal titration calorimetry/ITC, spectroscopy, and others. The full list is given in **Data S1** and was created by merging information from (Kastritis et al. 2011) and (Vreven et al. 2015). (For each complex, it reports the functional class, the experimental technique used to evaluate the binding affinity, the equilibrium dissociation constant (K_d), the Gibbs free energy (ΔG), the I-rmsd, the ΔASA.) In (Vangone & Bonvin 2015), a number of these experimental techniques were considered to be *non-reliable*: inhibition assays and fluorescence spectrophotometry methods, and all others for which only a few data points were reported (potentiometry, radioligand, reduction assay, and sedimentation, etc.). Here, as in (Vangone & Bonvin 2015), we define as *reliable*: ITC, SPR, spectroscopy and stopped-flow. Therefore, we defined the subset 'AffinityB2_rel' to be composed only by the 125 complexes measured with reliable experimental techniques. The corresponding subsets 'Kastritis_rel' and 'Vreven_rel' contain 98 and 27 complexes respectively. The whole AffinityB2_rel benchmark spans several biological functional classes: 24 antibody/antigens (A or AB with bound antibody), 22 enzyme/inhibitors (EI), 8 enzyme/substrates (ES), 8 enzyme/regulatory subunits (ER), 13 G-protein containing (OG), 17 membrane receptors (OR), and 33 miscellaneous (OX). (The proportion of these complexes is reported in **Figure 6A** and in **Supplemental Table S8**.) We also named 'Kastritis81', the subset of 81 complexes from AffinityB2_rel that was used in (Vangone & Bonvin 2015). For our analysis, we downloaded the 125 X-ray complex structures from the Protein Data Bank (Berman et al. 2002) without refining them nor discarding the hetero-atoms.

For each complex, experimental binding energy and I-rmsd (Interface rmsd; calculated in (Vreven et al. 2015) after superimposition of the unbound component proteins onto their bound forms, using the Cα atoms of the residues that had any atom within 10Å of any atom of the binding partner) were taken from (Vreven et al. 2015), and the ΔrASA (change in relative accessible surface area, rASA, upon complex formation (Kastritis et al. 2011)) was computed. Note that the values of binding energy vary in the range between -18.58 and -4.29 kcal/mol (Kastritis et al. 2011, Vreven et al. 2015). I-rmsd, which provides an estimation of the amplitude of the conformational changes that take place upon binding (Vangone & Bonvin 2015), varies between 0.17Å and 4.9Å. Moreover, in the context of affinity prediction, the complexes with I-rmsd $<$ 1Å were considered as rigid, those corresponding to 1Å $\leq$ I-rmsd $\leq$ 2Å as medium-rigid, and the remaining ones as flexible (Vreven et al. 2012). We also considered a coarser classification of the complexes into rigid (I-rmsd $\leq$ 1Å) and flexible (I-rmsd $>$ 1Å) as in (Vangone & Bonvin 2015).

Interface residues. The relative accessible surface area (rASA) was computed with NACCESS (Hubbard & Thornton 1993) using a probe radius of 1.4Å. According to (Levy 2010), the interface is identified by the residues that lose rASA upon binding. By computing rASA for the (bound form of the) monomers and for the complex, we define residues lying at the interface of a complex as those whose rASA in the complex is different than that in the monomer (ΔrASA$\geq$ 0).

Support-Core-Rim model of protein interfaces. Interface residues were subdivided in three classes according to their position at the interface (Levy 2010): 'support' residues are buried in unbound and bound forms ($rASA_u$, $rASA_b < 0.25$), 'core' residues become buried upon binding to the partner

($\text{rASA}_\text{u} \geq 0.25$ and $\text{rASA}_\text{b} < 0.25$), and 'rim' residues are exposed in unbound and bound forms (rASA_u, $\text{rASA}_\text{b} \geq 0.25$). We refer to this classification as the 'SCR model'. For protein-protein interactions these three regions of the protein interface roughly correspond to three concentric layers, the support defining the internal layer and the rim the most external one (Levy 2010, Laine & Carbone 2015).

Interface atom-atom contacts. Two residues, respectively belonging to the two proteins in a complex, are in contact with each other if the distance between their two nearest atoms is ≤ 5.5Å as in (Vangone & Bonvin 2015).

Favorable and non-favorable contacts. To classify a contact as *favorable* or *non-favorable* to the interaction, we use the quantum mechanical electron density, ρ, the key quantity in density functional theory (DFT), and two scalar fields derived from it, the reduced gradient (RDG, s) and the signed electron density (ϱ). With these measures, one can explore a wide range of non-covalent interaction types (van der Waals interactions, hydrogen-bonds and steric clashes) as surfaces. These measures are fundamental to the Non-Covalent Interaction (NCI) analysis approach (Johnson et al. 2010, Contreras-Garcia et al. 2011, Gunther et al. 2014), one of the three main computational modules in LISA.

The electron density ρ describes the relative probability to find electrons in a particular location of a space Ω. Density values are indicators of the interaction strength.

The reduced density gradient (RDG) s describes the deviation in atomic densities due to non-covalent interactions (Johnson et al. 2010, Contreras-Garcia et al. 2011, Gunther et al. 2014). It is defined as

$$s = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho|}{\rho^{4/3}} \tag{1}$$

Properties of s have been investigated in the process of developing increasingly accurate functionals (Hohenberg & Kohn 1964, Becke 1995, Cohen et al. 2008). In the presence of non-covalent interactions, s reports a strong change in its values in regions of space between interacting atoms. Namely, s shows large positive values in regions far from other molecules (monomers), in which the density is decaying to zero exponentially. Conversely, s assumes very small values, approaching zero, for regions of both covalent bonding and non-covalent interactions (upon dimer/complex formation).

In (Johnson et al. 2010), it was observed that plotting s (RDG) versus ρ (density) reveals the basic pattern of intra-molecular interactions. Visually, in plots of s versus ρ , "spikes" in the low-density low-gradient region are indicative of non-covalent interactions. The topological analysis of these signals reveals interaction sites which enable to focus on locations of space which are relevant for molecular interactions in chemical systems (Gunther et al. 2014).

Very different types of interactions might appear in the same region of density space and to distinguish between these them we consider the second eigenvalue, λ_2, of the electron density Hessian matrix. See (Johnson et al. 2010). λ_2 can be either positive or negative depending on the type of interaction: while attractive interactions concentrate electron charge perpendicular to the bond ($\lambda_2 \leq 0$), repulsive interactions cause density depletion ($\lambda_2 > 0$). In contrast to ρ which only assesses the interaction strength of atoms, the signed electron density, $\varrho = sign(\lambda_2) \times \rho$, additionally enables the differentiation of attracting and repulsive interactions. Both van der Waals interactions and hydrogen bonds show negative values of λ_2 at critical points (Johnson et al. 2010, Contreras-Garcia et al. 2011, Gunther et al. 2014). Namely, we define favorable/attractive interactions as non-covalent interaction contacts with $s \leq 1$ and $\varrho \leq 0$, and non-favorable/repulsive interactions as non-covalent interaction contacts with $s \leq 1$ and $\varrho > 0$.

In LISA, we use NCIPLOT v3 (Johnson et al. 2010) (`http://www.lct.jussieu.fr/pagesperso/contrera/nciplot.htm`), a program that enables the computation and graphical visualization of inter-

and intra-molecular non-covalent interactions (hydrogen bonds, π-π interactions, Van der Waals forces). It has been previously applied to systems ranging from small dimers to large biomolecules (Contreras-Garcia et al. 2011, Johnson et al. 2010). NCIPLOT performs the NCI (Non-Covalent Interaction) analysis, based on ρ, s, ϱ and λ_2. In our analysis, we concentrate only on the interactions concerning complex interfaces. For this, we apply NCIPLOT (with the "INCREMENTS" option to set the increments along the x,y,z directions of the cube - x=y=z=0.45 atomic units - and the "INTERMOLECULAR" option to analyze only inter-molecular interactions) to the NCI cube enveloping the protein-protein interface. To restrict the automatically generated NCI cube into a NCI cube that perfectly envelops the interface, we use the minimum and maximum atomic coordinates at the interface as boundary points (NCIPLOT option `cube`). An extra radial threshold in each direction was added to ensure that the interface atoms were contained within the cube (2 Å). A NCI grid, describing the interaction isosurface between the two monomers, was then generated (see **Figure 1A**) and the ρ, s, ϱ and λ_2 values were calculated at each point on the grid. Note that densities were obtained from promolecular estimates (ρ^{pro}) defined as the sum of atomic densities ($\rho_{\text{i}}{}^{\text{at}}$) $\rho^{\text{pro}} = \sum_i \rho_{\text{i}}{}^{\text{at}}$, where i varies through all atoms in the NCI grid. Promolecular densities are computationally very useful for describing large biomolecular systems (Contreras-Garcia et al. 2011).

In LISA, we use NCIPLOT in two different manners. We distinguish favorable from non-favorable non-covalent interactions on isosurfaces, and we exploit the three-dimensional localisation of molecular interactions identified in NCIPLOT by recording ρ, s, ϱ and λ_2 for each point in the NCI grid (**Figure 1A**). To each atom-atom contact at the protein complex interface, we associate the values ρ, s, ϱ and λ_2 of the nearest point in the NCI grid. This allows to determine the position of a contact in the NCI grid, its strength, and whether it is favorable or not.

Interface geometry and atom-atom contacts. To analyse the distribution of interface atom-atom contacts and their properties in 3D space, we define, for each complex, a cuboid grid enveloping all residues involved in the interfaces of the two molecules. (Note that this grid is not the NCI cubic grid.) To ensure that all interface interactions are included, we used the minimum (X^{at}_{min}, Y^{at}_{min}, Z^{at}_{min}) and the maximum (X^{at}_{max}, Y^{at}_{max}, Z^{at}_{max}) atomic coordinates at the interface. Then, the cube was subdivided in 125 cubic units of identical volume by dividing each axis in five parts such that each edge along the X-, Y- and Z-axis respectively, has size $l_x = \frac{X^{at}_{max} - X^{at}_{min}}{5}$, $l_y = \frac{Y^{at}_{max} - Y^{at}_{min}}{5}$, $l_z = \frac{Z^{at}_{max} - Z^{at}_{min}}{5}$, respectively. For each unit cube, the coordinates of a contact are those coming from the NCI cubic grid. We computed the total number of atom-atom contacts, the number of favorable and non-favorable contacts, and the total strength (provided by the absolute value of ϱ) of contacts in the cube. Following the SCR model (Levy 2010) (see above), we also classified the contacts in six classes depending on whether they belong to RR: rim-rim, CC: core-core, SS: support-support, RS: rim-support, RC: rim-core, and CS: core-support. Strictly speaking, residues are classified as belonging to rim, core or support, and here, we consider the obvious extension to atoms.

Each unit cube belongs to one of the three families, IS_1, IS_2 and IS_3, representing the intensity of the favorable strength (IS) in the unit cube. Given the total strength of favorable contacts in each cube, we consider their distribution and rescale it on the interval $[0, 1]$. Then, we define the families IS_1, IS_2 and IS_3 in such a way that IS_1 contains all cubes with normalised favorable strength ≥ 0.6, IS_2 contains all cubes with normalised favorable strength between 0.1 (included) and 0.6 and IS_3 contains all cubes with normalised favorable strength < 0.1. Note that normalisation is used only to classify the cubes in the families.

Structural properties and NIS at the interface. Non-interacting surface (NIS) contribution in protein-protein binding (Kastritis et al. 2014) was analysed by considering distinguished physico-chemical properties of residues lying on the complex surface: polar (C, H, N, Q, S, T, Y, W; $\text{NIS}_{\text{polar}}$), apolar (A, F, G, I, L, V, M, P; $\text{NIS}_{\text{apolar}}$) and charged (E, D, K, R; $\text{NIS}_{\text{charged}}$). Calculations for $\text{NIS}_{\text{polar}}$, $\text{NIS}_{\text{apolar}}$ and $\text{NIS}_{\text{charged}}$ were realised with scripts available at `http://bonvinlab.org/software`. We ran them with default options and default parameter values.

The secondary structure at the interface was computed by using the DSSP program (Kabsch & Sander 1983*b*, Joosten et al. 2011), the standard software for the assignment of secondary structure elements in PDB entries. We exploited the full classification of structural elements used in DSSP: α-helices, 3_{10}-helices, π-helices, B-bridges, β-sheets, turns, bends, loops (Kabsch & Sander 1983*a*). DSSP was used with default options and default parameter values; it is accessible at `http://swift.cmbi.ru.nl/gv/dssp/`.

Local Interaction Signal Analysis (LISA) and the LISA score. LISA aims at determining geometrical features of atom-atom contact interactions that effectively influence binding energy. To construct LISA model, we started with 179 descriptors defined by combining properties associated to either the LISA cube enveloping the interface or the NCI cubic grid. When we consider the LISA cube enveloping the interface, we highlight 6 main properties: the total number of contacts, the total number of favorable contacts, the total number of non-favorable contacts, the total strength of the contacts, the strength obtained by considering only favorable contacts and the strength obtained from non-favorable contacts. These properties are studied by restricting the contacts to those belonging to SS, CC, RR, SR, SC and CR residue pairs. This provides 36 more descriptors. When we consider the NCI cubic grid, we combine the 42 properties with the three families IS_1, IS_2 and IS_3. This makes 126 other descriptors. We also consider 8 parameters describing the interaction surfaces of both receptor and ligand as secondary structures. Namely, we count the number of residues in the interface that belong to α-helices, 3_{10}-helices, π-helices, B-bridges, β-sheets, turns, bends, loops. The 3 last descriptors are NIS_{polar}, NIS_{apolar} and NIS_{charged}, describing the proportion of polar (C, H, N, Q, S, T, Y, W), apolar (A, F, G, I, L, V, M, P) and charged (E, D, K, R) residues that lie at the complex surface and that do not belong to the complex interface. Notice that all parameters collect values coming from the bound structure. In **Data S2**, for each complex, we report the values of the 179 features used to evaluate the models and compute the LISA score of the complexes.

We used the Akaike's Information Criterion (AIC) feature stepwise selection approach (backward and forward) in order to choose a linear regression model identifying the significant features and to avoid over-training problems due to the presence of more than three variables. The AIC-based linear regression stepwise procedure used the Kastritis_rel dataset as a training set. The chosen model contains 58 features, associated with coefficients and p-values. The 58 features are then ordered by p-values and the ensemble is progressively reduced by applying the following procedure: for each feature f_e, we remove the features f_i that have a p-value $pval(f_i) > pval(f_e)$ and that are highly correlated with f_e ($R > 0.75$). We end up with 21 features. From this ensemble, we remove features with a p-value > 0.1 or a number of null values > 60 (more than two thirds of the complex set). We end up with 16 features, which are used to perform a Support Vector Regression (SVR).

Namely, we want to select the best combination of features among the 16 ones, by running SVR based on a radial kernel, on the Kastritis_rel dataset used as a training set. The feature's selection is realized by running calculations for all 65 535 combinations of the 16 features on the training set. Then, each model is used to predict binding affinities for the Vreven_rel dataset, that is the testing set, and the correlations between scores produced by the model and the experimental binding affinities are

computed on both Kastritis_rel and Vreven_rel datasets with the Pearson product-moment correlation coefficient, a widely accepted measure of the relationship between predicted and experimental binding energy (Moal et al. 2011). The best combination of features is chosen as the one that maximizes the sum of the correlations obtained with the training set (Kastritis_rel) and the testing set (Vreven_rel) over all models. The list of selected features (V46, V69, V106, V107, V114, V154, V202, V207, V208 and nis2) is described in **Supplemental Table S1** and **Supplemental Figure S3**. Note that the list of features obtained by optimizing only on Kastritis_rel (V43, V46, V69, V97, V107, V123, V202, V207, V208 and nis2) provides correlation coefficients of 0.87 for Kastristis_rel and 0.43 for Vreven_rel. Compare it with 0.84 for Kastristis_rel and 0.67 for Vreven_rel obtained by LISA model. We stress that our feature selection procedure is completely automatized and much more exhaustive than what is usually done for the development of empirical scoring functions. In particular, we have built 65 535 regression models corresponding to all possible combinations of our 16 pre-selected features. By comparison, for example in (Vangone & Bonvin 2015), only 6 models were built and the choice of the features in each model was largely manually driven.

For each combination of n features, to find a radial function of support vectors of dimension n, we used the SVR included in the package e1071 in R (R Development Core Team 2011), with a radial kernel, default "cost of constraints violation" and default "epsilon in the insensitive-loss function". The kernel function $K(x, x') = exp(-\gamma \parallel x - x' \parallel^2)$ (where $\gamma = 0.1$ by default) is optimized to extract N representative support vectors that are used for classification. The *LISA score* is the classification function defined as $\sum_{i=1}^{N} \alpha_i K(x_i, x') + b \geq 0$, where the parameters α_i are learned in the SVR training step over the 10 selected features. Notice that there are as many α_i parameters as support vectors, in our case $N = 84$.

The SVR model on 10 features is validated through 100 cycles of 4-fold cross-validation helping to estimate the Root Mean Square Error (RMSE) between predicted and experimental binding affinity. The procedure goes as follows: Kastritis_rel is split into four equal subsets and we chose one of them as a testing set, while the other three datasets are used to run a SVR with a radial kernel. This allows us to find the model best fitting the 75% of the data. Then, this model is used to predict scores for complexes in the testing set and compute the RMSE between these values and the experimental binding affinity values. This procedure is repeated for the 4 distinguished subsets and for 100 times on different random splittings of Kastritis_rel. Hence, we generated 4×100 models with different parameters and computed 4×100 RMSE values, whose distribution allowed us to determine the mean and the standard deviation representing the quality of the model.

Procedure for the evaluation of LISA modules. For each discarded LISA module, we re-compute the model anew, starting from those features that characterise the modules. The flow of the computation remains the same as for LISA, while the set of initial parameters is smaller. The analyses are realised on the Kastritis_rel (used for training) and Vreven_rel (used for testing) datasets. The RMSE has been computed for each model on the Kastritis_rel dataset with a 4-fold cross-validation procedure.

Comparison with other tools. We compared LISA with other 17 available tools. In **Data S3**, for each complex, we report the LISA score and the Gibbs free energy (ΔG). ZAPP (Vreven et al. 2012) and ConsBind (Moal et al. 2011) correlations were collected from (Vreven et al. 2012, 2015) and (Moal et al. 2011, Vreven et al. 2015), respectively. Prodigy (Vangone & Bonvin 2015) data were calculated through the program used with default options and default parameter values. All the results related to other methods used for comparison were obtained as pre-computed data from the CCharPPI server (Moal et al. 2015) (`https://life.bsc.es/pid/ccharppi/`) using external tools in order to calculate

different descriptors. The descriptors we considered are classified as composite scoring functions (FireDock, FireDock_EI, FireDock_AB (Andrusier et al. 2007), ZRANK (Pierce & Weng 2007), ZRANK2 (Pierce & Weng 2008), RosettaDock (Chaudhury et al. 2010), PyDoc_Tot (Grosdidier et al. 2007, Jiménez-García et al. 2013), Sipper (Pons et al. 2011), AP_Pisa (Viswanath et al. 2013), CP_PIE (Ravikant & Elber 2010)), atomic-distance dependent potentials (AP_DComplex (Liu et al. 2004), AP_dDFire (Yang & Zhou 2008*b*), AP_dDFire2 (Yang & Zhou 2008*a*)) and miscellaneous (INSIDE (Andrusier et al. 2007)).

The two tools, GA-PLS (Tian et al. 2012) and BioQSAR (Biomacromolecular Quantitative Structure-Activity Relationship) (Zhou et al. 2013), both based on a QSAR model, could not be considered for comparison because of software unavailability. Direct enquiry to the authors remained with no answer. It should be noticed that GA-PLS claims a correlation coefficient of 0.83 and BioQSAR of 0.82-0.88 (with RMSE of 0.8-1.5 kcal/mol) on the 144 complexes of the AffinityB2 dataset, where both tools have been trained and tested. Since experimental binding affinities for AffinityB2 complexes were measured by reliable and unreliable experimental techniques, contrary to our choice of considering only reliable techniques as in (Vangone & Bonvin 2015), we could not directly compare LISA's performance to GA-PLS and BioQSAR.

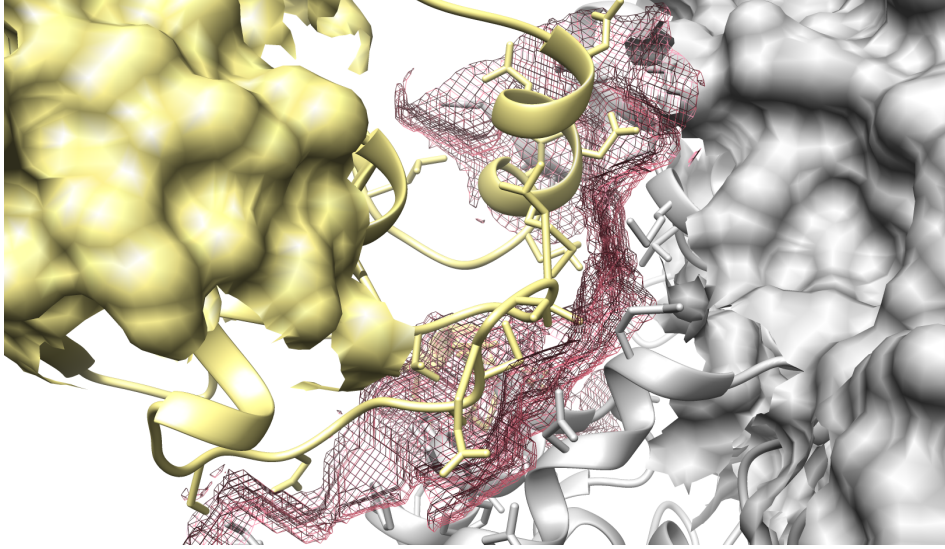
Quantification and statistical analysis. The development of the LISA package is based on Support Vector Regression, as described above. Comparison with other tools is based on Pearson correlation coefficients computed with the R package (R Development Core Team 2011).

Software availability. LISA package is freely available for the community at `http://www.lcqb.upmc.fr/LISA/` under the CeCILL licence.

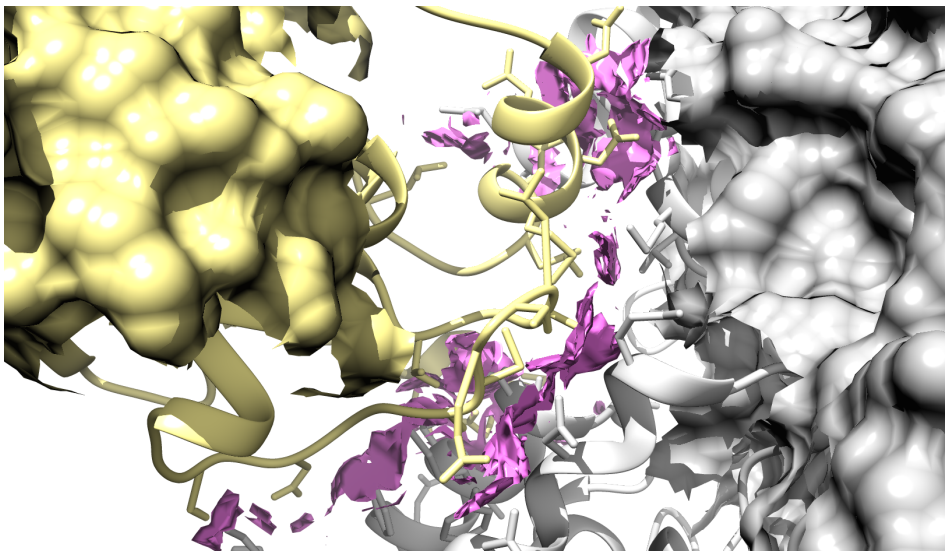
Description of supplemental files

Supplemental Figures and Tables: supplemental figures and supplemental tables.
Data S1, related to Figures 4 and 6: for each complex, we report the functional class, the experimental technique used to evaluate the binding affinity, the equilibrium dissociation constant (K_d), the Gibbs free energy (ΔG), the I-rmsd, the ΔASA.
Data S2, related to Figures 2, 3, 4 and Table 1: for each complex, we report the values of the 179 features used to evaluate the models and compute the LISA score of the complexes.
Data S3, related to Figure 4: for each complex, we report the LISA score and the Gibbs free energy (ΔG).

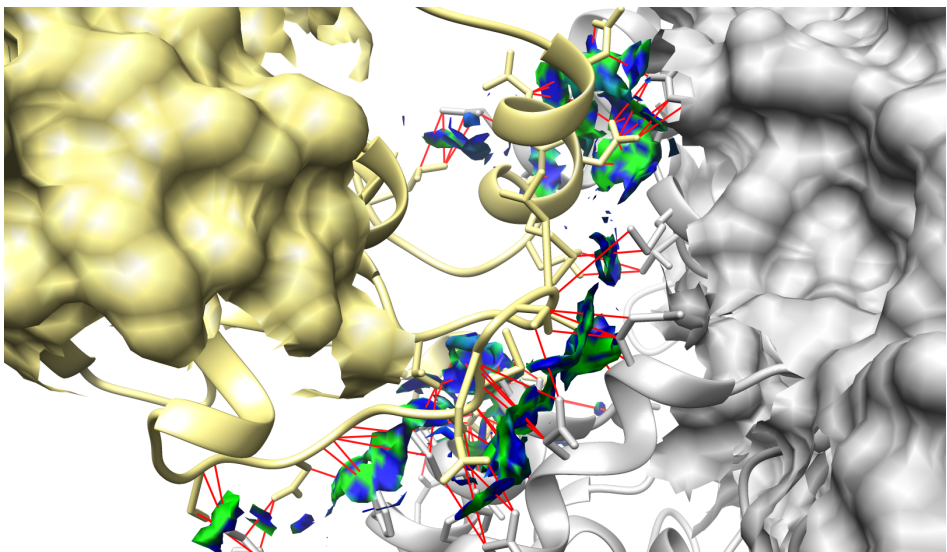
A.



B.



C.



NCI cube

NCI grid

NCI isosurfaces

Atom-atom contacts

NCI detection of favorable and non-favorable contacts

- atom-atom contacts
- favorable contacts (RDG ≤ 1, ϱ < 0)
- non-favorable contacts (RDG ≤ 1, ϱ ≥ 0)

SCR model for protein interfaces

LISA cuboid grid

- Support
- Core
- Rim

DSSP secondary structure analysis

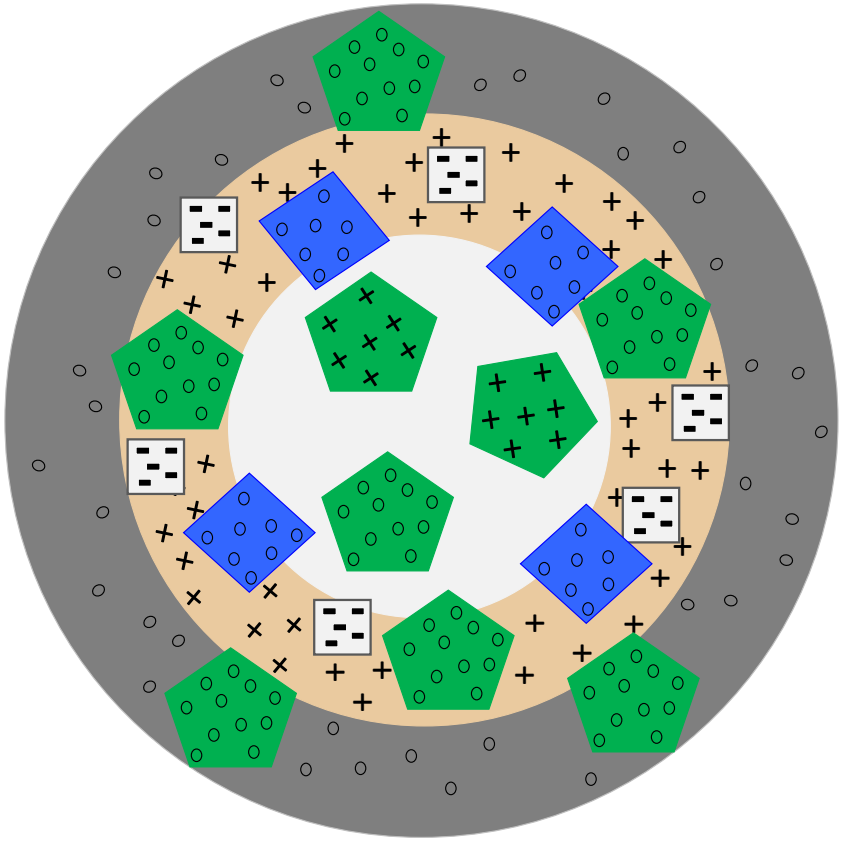
- Beta-sheet
- 3_{10}-helix
- Bend
- Turn
- Alpha-helix

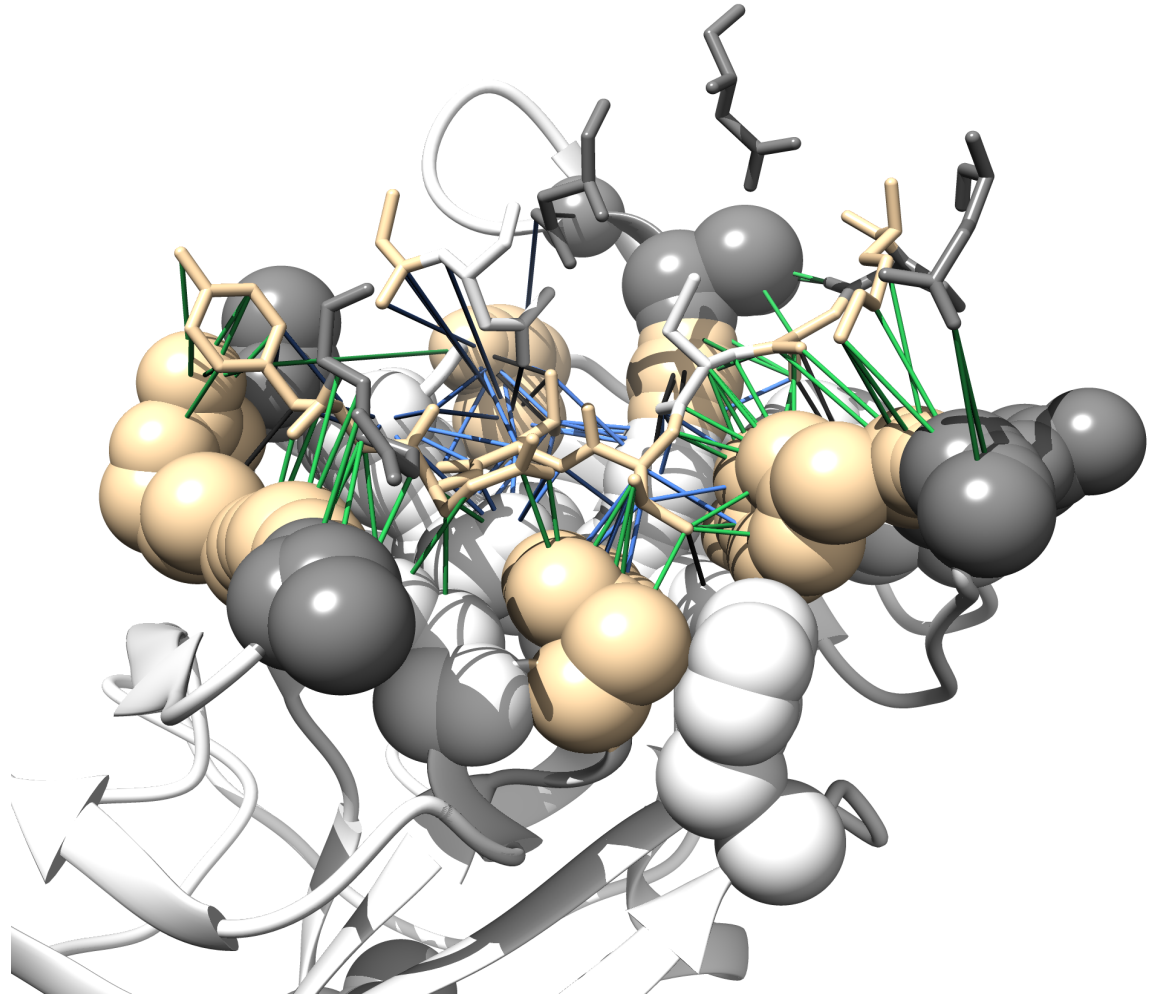
LISA scoring

NIS analyses

- Charged residues

- favorable contacts in high strength regions
- favorable contacts in medium strength regions
- unfavorable contacts in low strength regions





High strength regions

CC/SS/RS Medium strength regions

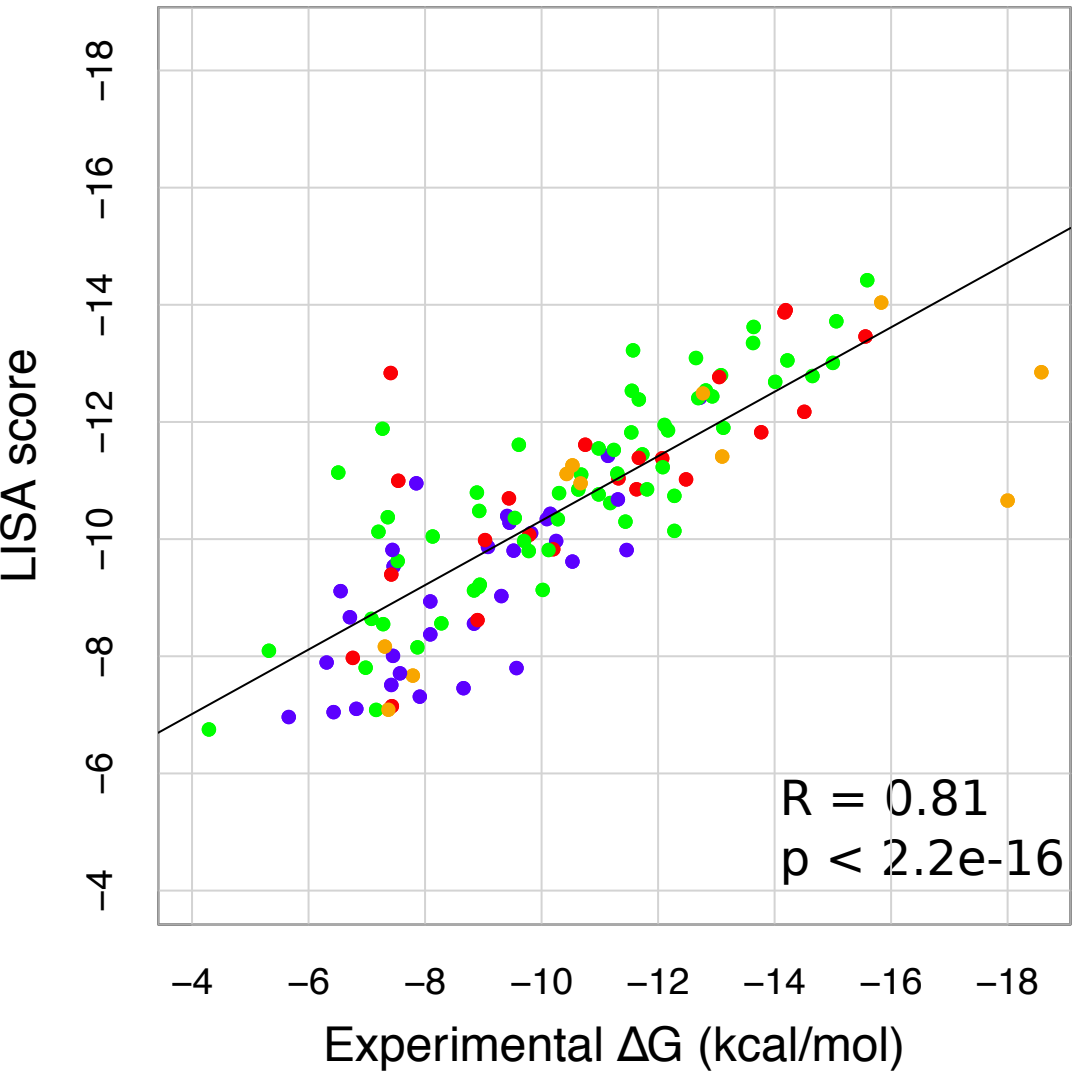
\+ Contacts

° Favorable contacts

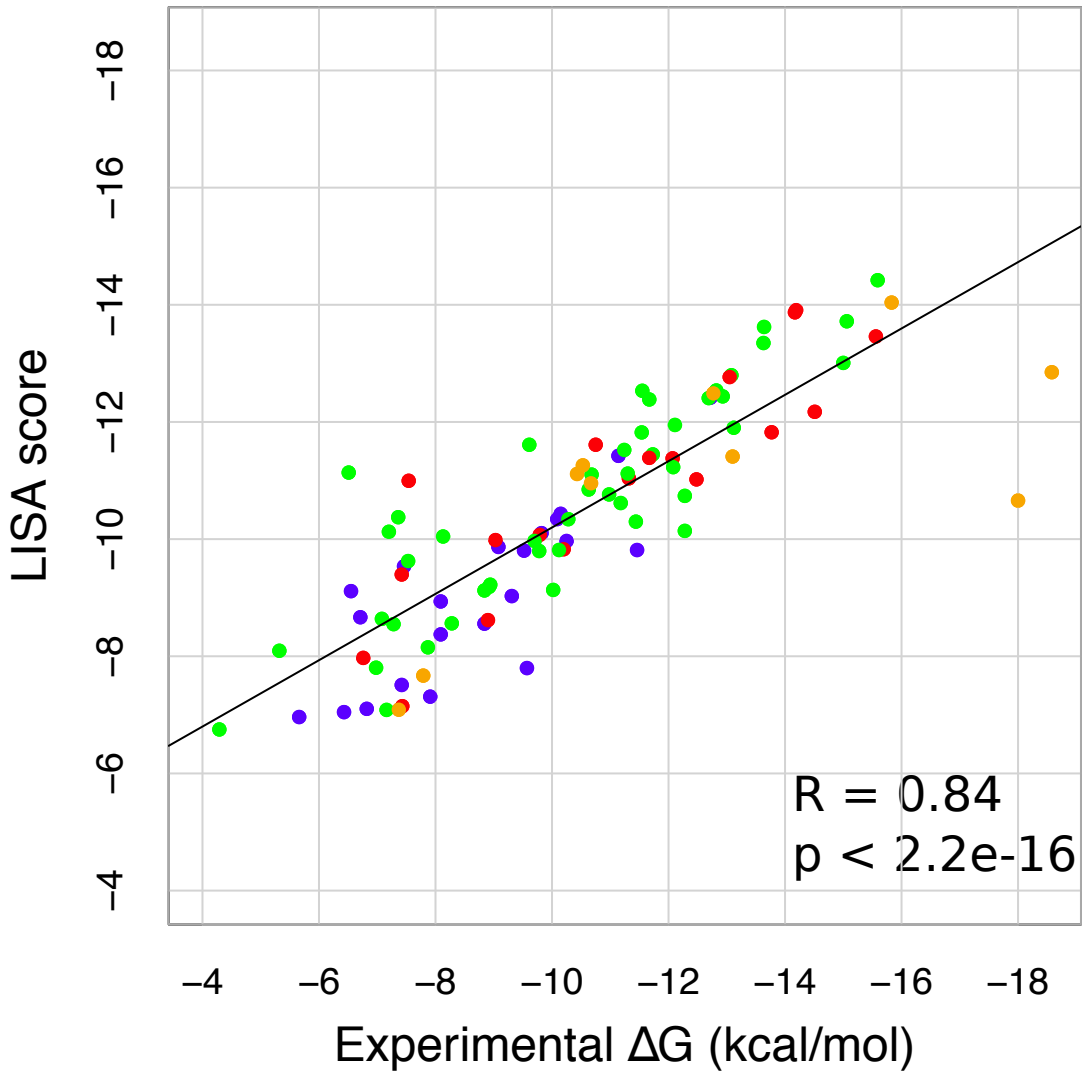
Low strength unfavorable regions

▪ Unfavorable contacts

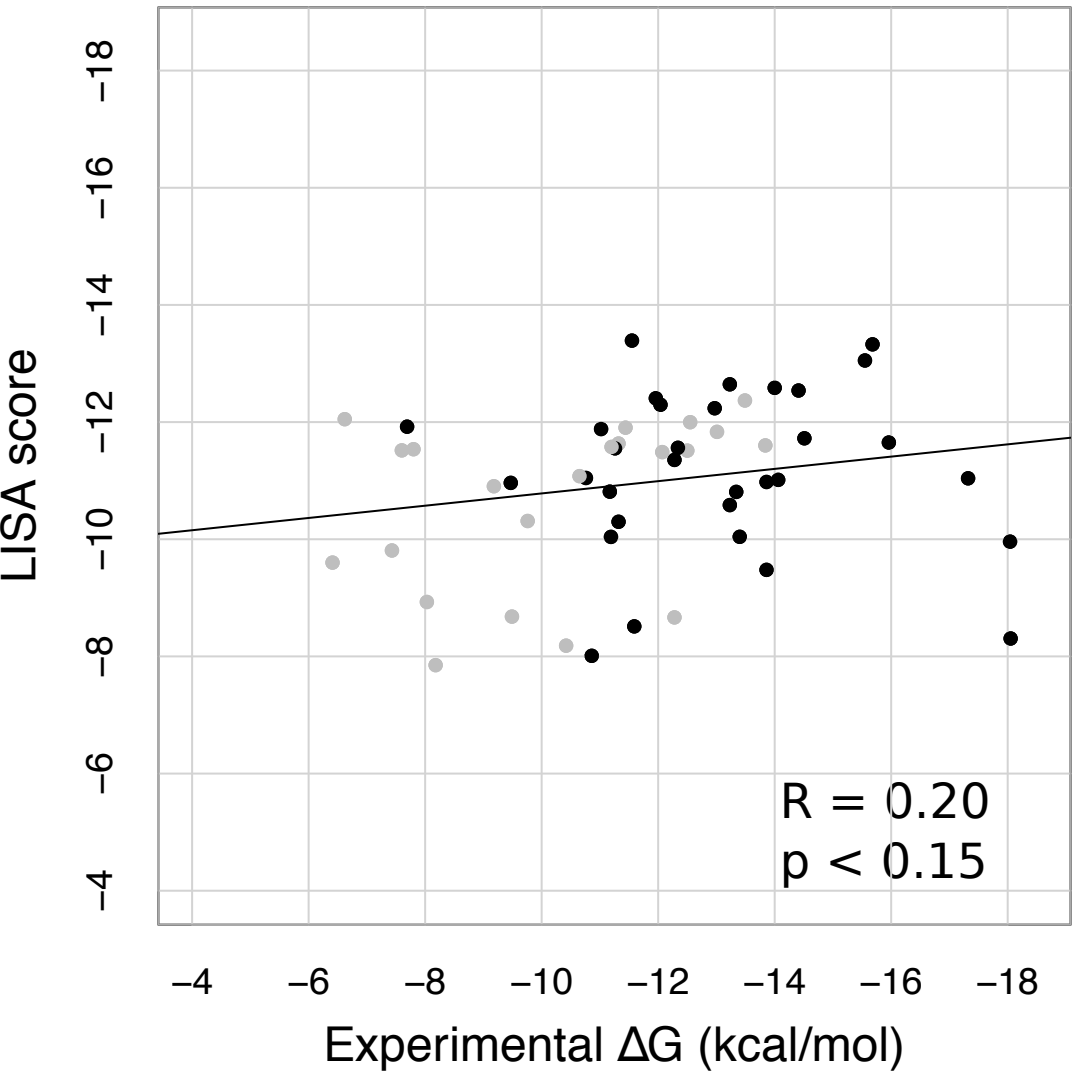
Affinity B2 - reliable experimental methods



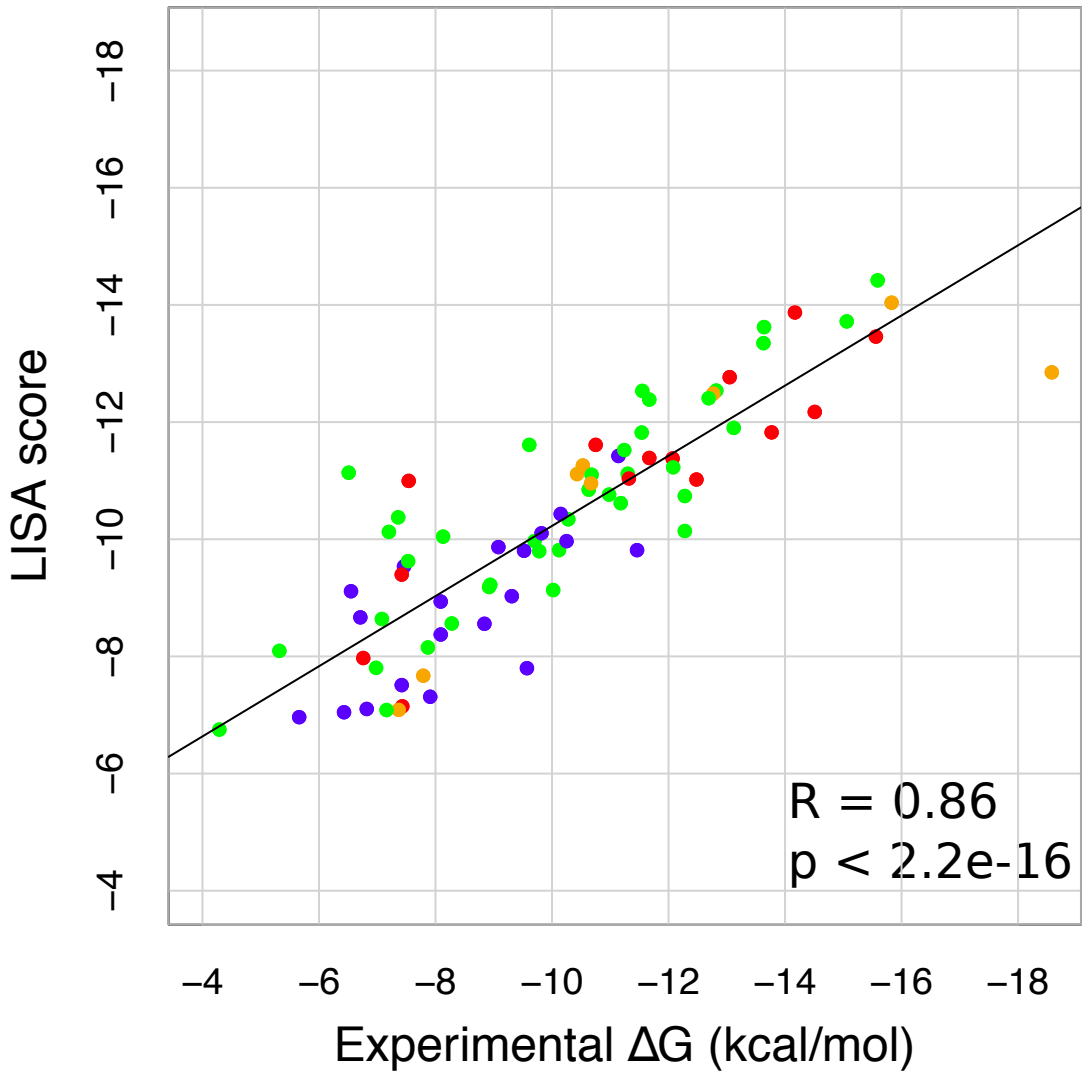
Kastritis - reliable experimental methods



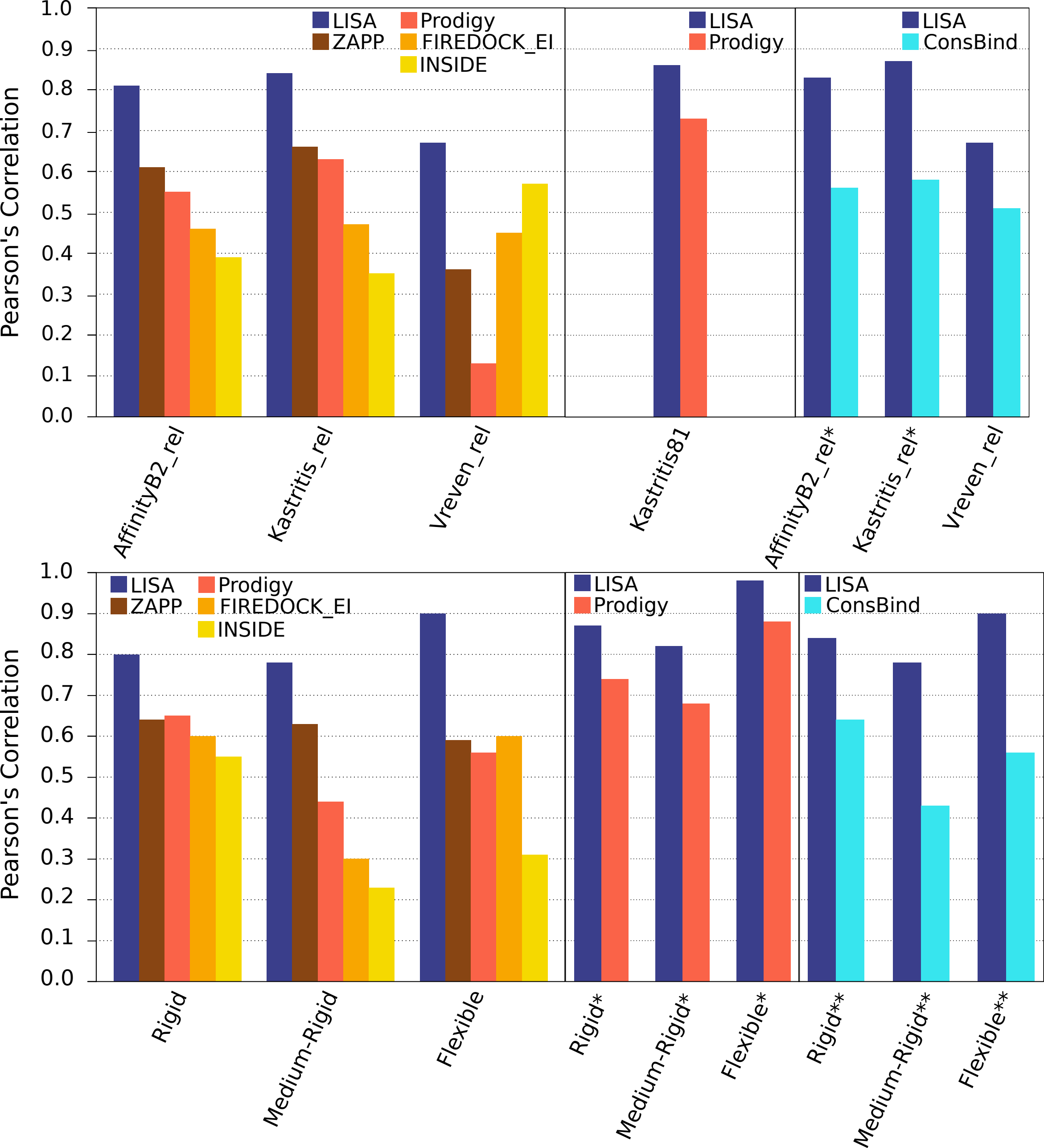
Affinity B2 - unreliable experimental methods

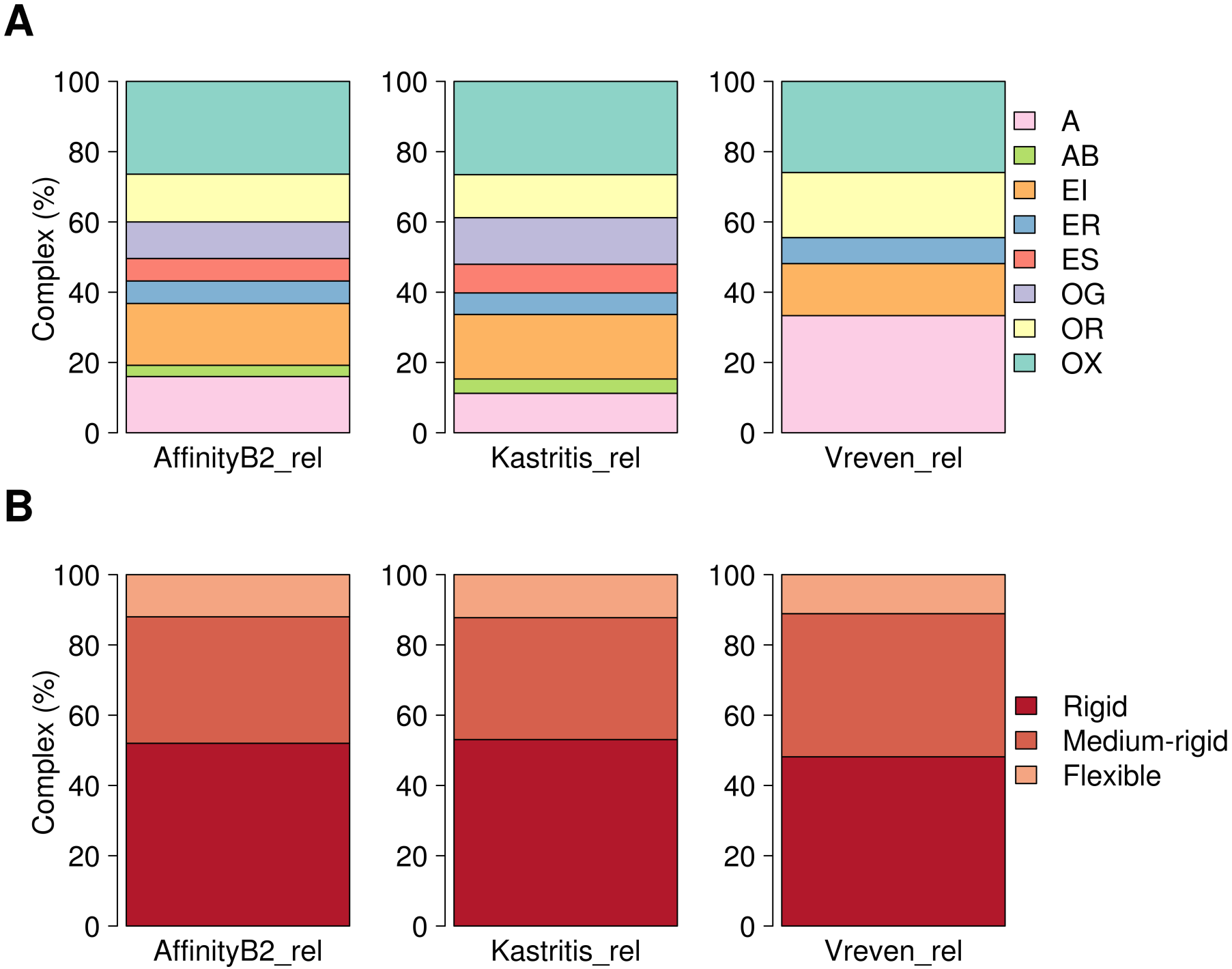


Kastritis81



Inhibition assay ITC Others Spectroscopy SPR Stopped-flow


Pearson's Correlation
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
LISA
ZAPP
Prodigy
FIREDOCK_EI
INSIDE
LISA
Prodigy
LISA
ConsBind
AffinityB2_rel
Kastritis_rel
Vreven_rel
Kastritis81
AffinityB2_rel*
Kastritis_rel*
Vreven_rel
Rigid
Medium-Rigid
Flexible
Rigid*
Medium-Rigid*
Flexible*
Rigid**
Medium-Rigid**
Flexible**


A
100
80
60
40
20
0
Complex (%)
AffinityB2_rel
Kastritis_rel
Vreven_rel
A
AB
EI
ER
ES
OG
OR
OX
B
Rigid
Medium-rigid
Flexible

SUPPLEMENTARY FIGURES

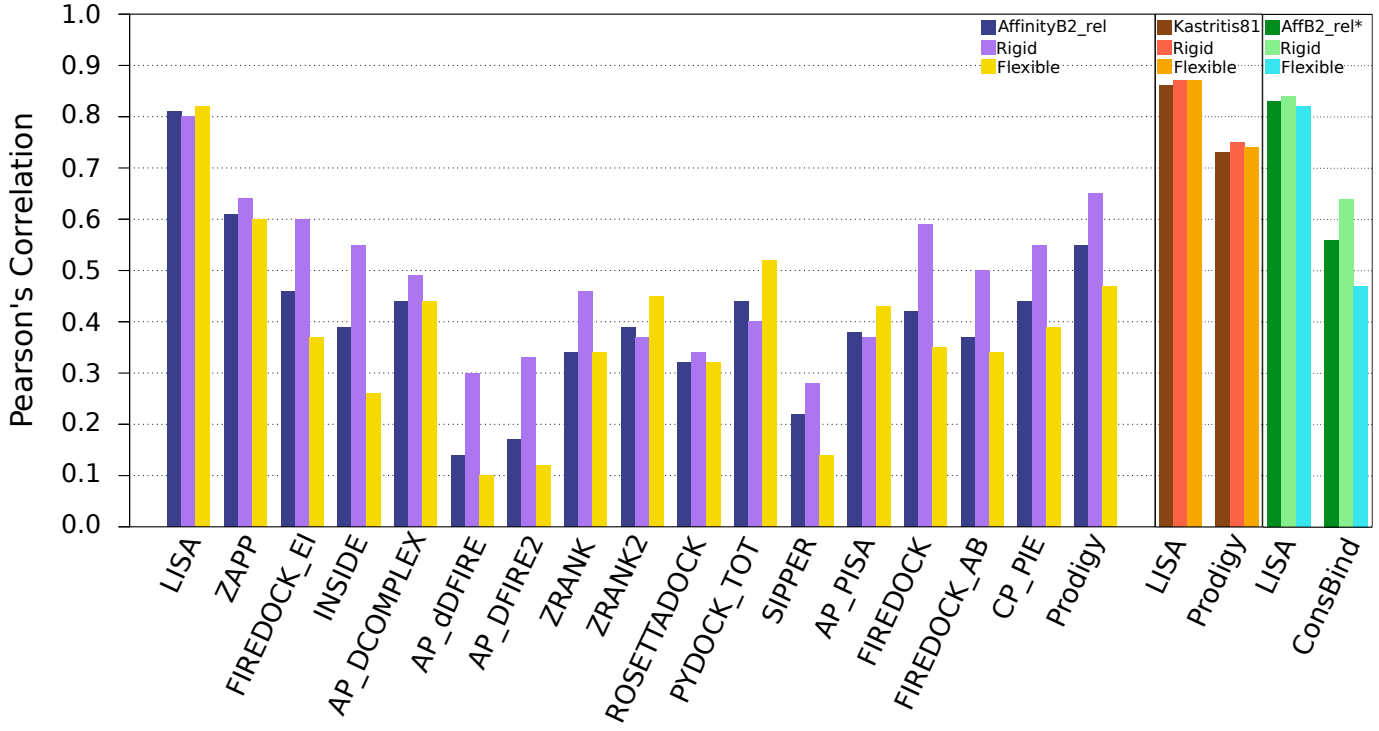


Figure S1. LISA performance and comparison with 17 other predictive tools. Performance of LISA and comparison with other tools on the 125 protein complexes in the AffinityB2_rel dataset. Each comparison provides correlations on AffinityB2_rel (blue) and on its two complementary subsets characterised by rigid (violet) and flexible (yellow) complexes (Vangone & Bonvin 2015). ConsBind is applied on the AffinityB2_rel* dataset and its comparison to LISA is reported on the right plot (green colorscale). In (Vangone & Bonvin 2015), Prodigy has been evaluated on the Kastritis81 dataset and, here, we report the corresponding comparison to LISA restricted to this subset (central plot; red colorscale); compare it with the evaluation on AffinityB2_rel, on the left. We recall that LISA is trained on the Kastritis_rel dataset. See **Table S7** for the numerical values of the correlations plotted on the left, **Table S8** for the plot on the right and **Table S9** for the plot on the center.

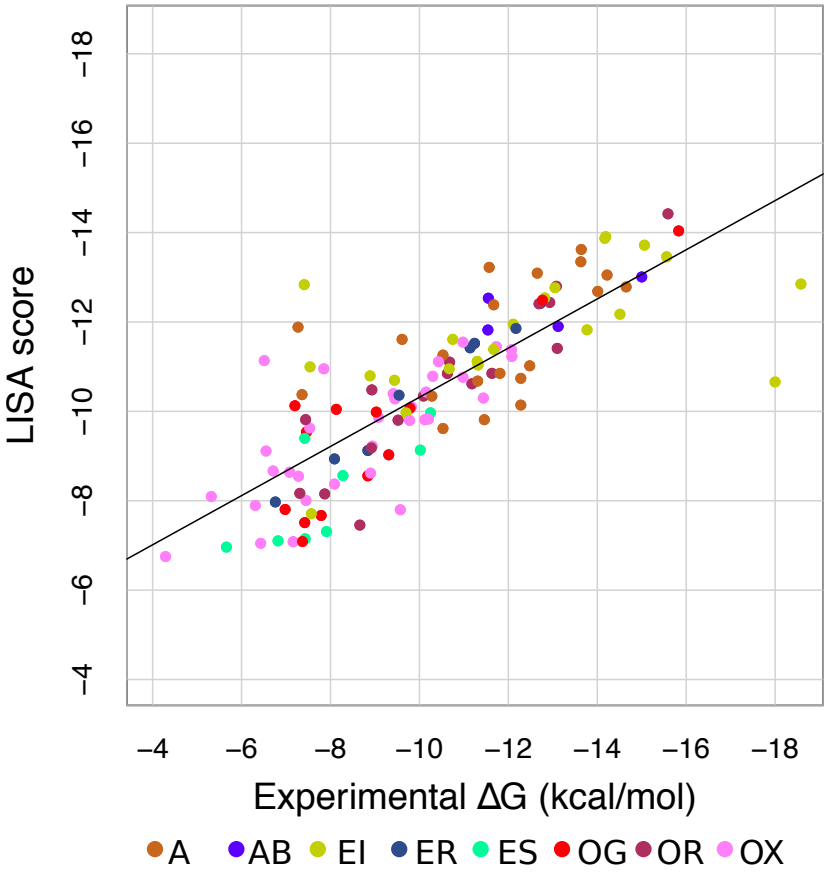


Figure S2. Correlation of LISA scores with binding affinities, organised by functional classes. Correlation (0.72) between the binding affinity predicted by LISA on the 125 protein complexes in AffinityB2_rel and their experimental binding energies reported in (Vreven et al. 2015). Complexes are coloured with respect to functional classes: antibodies (A), bound antibodies (AB), enzyme/inhibitors (EI), enzyme/substrate (ES), enzyme/regulatory subunit (ER), G-protein containing (OG), membrane receptors (OR), and miscellaneous (OX). Compare with **Figure 4** (top left).

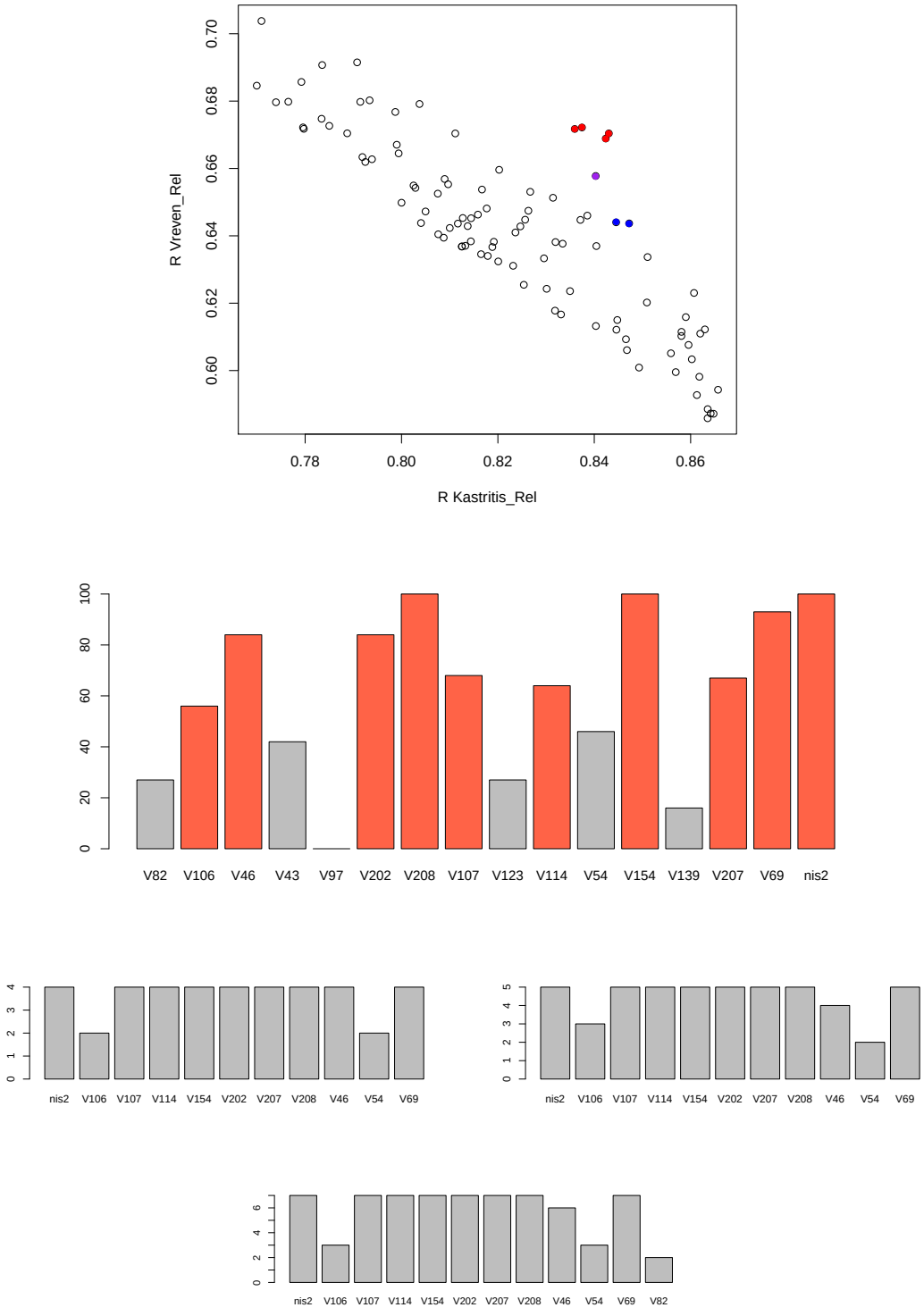


Figure S3. Analysis of LISA features for the 100 best scoring feature combinations. The 100 feature combinations providing the best sum of correlations on Kastritis_rel and Vreven_rel datasets are considered. **Top**: each point correspond to a feature combination. **Center**: each bin reports the number of times a feature belongs to the 100 best feature combinations. The 16 features selected by the AIC procedure are considered: $\text{TOT}_{\text{fav}}|IS2$ (V82), $\text{Strength}|CC|IS_2$ (V106), $\text{TOT}|CC$ (V46), $\text{Strength}_{\text{fav}}$ (V43), $\text{TOT}_{\text{fav}}|RC|IS2$ (V97; it does not occur in any of the best 100 feature combinations), $\text{Strength}_{\text{unfav}}|CC|IS_3$ (V202), 3_{10}-helix (V208), $\text{Strength}|SS|IS_2$ (V107), $\text{Tot}|IS1$ (V123), $\text{Strength}_{\text{fav}}|RS|IS_2$ (V114), $\text{TOT}_{\text{fav}}|RS$ (V54), $\text{Strength}_{\text{fav}}|CC|IS_1$ (V154), $\text{TOT}_{\text{fav}}|RC|IS_1$ (V139), α-helix (V207), $\text{Strength}_{\text{fav}}|RR$ (V69), $\text{NIS}_{\text{charged}}$ (nis2). Features are ordered with respect to the best Pearson correlation coefficient R between the feature values and the experimental binding affinities of the AffinityB2_rel complexes. Dark orange bins correspond to the 10 LISA features. **Bottom**: distribution of features in the first 4 (red dots in Top plot; top left), 5 (red and violet dots in Top plot; top right) and 7 (red, violet and blue dots in Top plot; bottom) best configurations.

SUPPLEMENTARY TABLES

Feature	Description	Coef.	R
Strength$\|CC\|IS_2$	Sum of the strengths of all CC contacts[1] in IS_2	-0.37	-0.48
Tot$\|CC$	Total number of CC contacts[1]	-0.22	-0.45
$\text{Strength}_{\text{unfav}}\|CC\|IS_3$	Sum of the strengths of all unfavorable CC contacts[1] in IS_3	-0.54	-0.29
3_{10}-helix	The existence of a 3_{10}-helix structure on the interaction surface	-0.18	-0.28
Strength$\|SS\|IS_2$	Sum of the strengths of all SS contacts[2] in IS_2	2.87	-0.17
$\text{Strength}_{\text{fav}}\|RS\|IS_2$	Sum of the strengths of all favorable RS contacts[3] in IS_2	0.77	-0.13
$\text{Strength}_{\text{fav}}\|CC\|IS_1$	Sum of the strengths of all favorable CC contacts[1] in IS_1	-1.45	-0.07
α-helix	The existence of a α-helix structure on the interaction surface	-0.1	0.02
$\text{Strength}_{\text{fav}}\|RR$	Sum of the strengths of all favorable RR contacts[4]	0.88	0.07
$\text{NIS}_{\text{charged}}$	Percentage of charged residues not involved in the direct interaction between two proteins in a complex	0.22	0.31

[1] Contacts between atoms belonging to core residues.
[2] Contacts between atoms belonging to support residues.
[3] Contacts between atoms belonging to rim and support residues.
[4] Contacts between atoms belonging to rim residues.

Table S1. Ten selected features characterising the LISA model. For each feature, we report the coefficient of the linear regression model computed from the AIC-based feature selection, and the Pearson correlation coefficient R between the feature values and the experimental binding affinities. Features are ordered with respect to R.

	AffinityB2_rel	Kastritis_rel	Vreven_rel
LISA	**0.81**	**0.84**	**0.67**
ZAPP	0.61	0.66	0.36
Prodigy	0.55	0.63	0.13
FIREDOCK_EI	0.46	0.47	0.45
INSIDE	0.39	0.35	0.57
———	—	—	—
AP_DCOMPLEX	0.44	0.46	0.27
AP_dDFIRE	0.13	0.14	0.14
AP_DFIRE2	0.16	0.17	0.16
ZRANK	0.33	0.34	0.4
ZRANK2	0.38	0.39	0.46
ROSETTADOCK	0.32	0.37	0.04
PYDOCK_TOT	0.44	0.46	0.32
SIPPER	0.22	0.18	0.46
AP_PISA	0.38	0.39	0.3
FIREDOCK	0.42	0.41	0.45
FIREDOCK_AB	0.37	0.37	0.37
CP_PIE	0.44	0.45	0.39

Table S2. Correlations on AffinityB2_rel, Kastritis_rel and Vreven_rel datasets. Correlation values between experimental and predicted binding energies for all complexes in the AffinityB2_rel, Kastritis_rel and Vreven_rel datasets.

I-rmsd range	AffinityB2_rel	I-rmsd < 1Å	1Å$\leq$ I-rmsd $\leq$ 2Å	I-rmsd > 2Å
LISA	**0.81**	**0.8**	**0.78**	**0.9**
ZAPP	0.61	0.64	0.63	0.59
Prodigy	0.55	0.65	0.44	0.56
FIREDOCK_EI	0.46	0.6	0.3	0.6
INSIDE	0.39	0.55	0.23	0.31
———	—	—	—	—
AP_DCOMPLEX	0.44	0.49	0.39	0.7
AP_dDFIRE	0.14	0.3	0.08	0.22
AP_DFIRE2	0.17	0.33	0.09	0.26
ZRANK	0.34	0.46	0.26	0.75
ZRANK2	0.39	0.37	0.37	0.7
ROSETTADOCK	0.32	0.34	0.23	0.64
PYDOCK_TOT	0.44	0.4	0.47	0.66
SIPPER	0.22	0.28	0.26	0.22
AP_PISA	0.38	0.37	0.35	0.72
FIREDOCK	0.42	0.59	0.24	0.74
FIREDOCK_AB	0.37	0.5	0.2	0.8
CP_PIE	0.44	0.55	0.35	0.53

Table S3. Correlations organised by I-rmsd ranges characterising rigid, medium-rigid and flexible complexes in AffinityB2_rel. Correlations between experimental binding energy and predictions obtained with several tools and organised by different I-rmsd intervals on AffinityB2_rel complexes. Recall that complexes with I-rmsd < 1Å, 1Å$\leq$ I-rmsd $\leq$ 2Å and I-rmsd > 2Å are classified as rigid, medium-rigid and flexible.

Intersection	Kastritis81	I-rmsd < 1Å	1Å$\leq$ I-rmsd $\leq$ 2Å	I-rmsd $>$ 2Å
LISA	**0.86**	**0.87**	**0.82**	**0.98**
Prodigy	0.73	0.74	0.68	0.88

Table S4. Correlations organised by I-rmsd ranges characterizing rigid, medium-rigid and flexible complexes in Kastritis81. Correlations between experimental and predicted binding energies obtained with LISA and Prodigy and organised by different I-rmsd intervals on Kastritis81 complexes. Recall that complexes with I-rmsd < 1Å, 1Å$\leq$ I-rmsd $\leq$ 2Å and I-rmsd $>$ 2Å are classified as rigid, medium-rigid and flexible in (Vreven et al. 2012).

	AffinityB2_rel*	Kastritis_rel*	Vreven_rel
LISA	**0.83**	**0.87**	**0.67**
ZAPP	0.62	0.67	0.36
Prodigy	0.54	0.62	0.13
FIREDOCK_EI	0.46	0.47	0.45
INSIDE	0.4	0.36	0.57
ConsBind	0.56	0.58	0.51
————————	—	—	—
AP_DCOMPLEX	0.45	0.48	0.27
AP_dDFIRE	0.14	0.13	0.14
AP_DFIRE2	0.17	0.16	0.16
ZRANK	0.33	0.33	0.4
ZRANK2	0.4	0.39	0.46
ROSETTADOCK	0.32	0.36	0.04
PYDOCK_TOT	0.42	0.44	0.32
SIPPER	0.21	0.16	0.46
AP_PISA	0.35	0.36	0.3
FIREDOCK	0.41	0.4	0.45
FIREDOCK_AB	0.36	0.36	0.37
CP_PIE	0.44	0.45	0.39

Table S5. Correlations on AffinityB2_rel, Kastritis_rel and Vreven_rel complexes belonging to the ConsBind dataset. Correlation values between experimental and predicted binding energies for all complexes in the intersection between AffinityB2_rel, Kastritis_rel, Vreven_rel and ConsBind datasets.

I-rmsd range	AffinityB2_rel*	I-rmsd < 1Å	1Å≤ I-rmsd ≤ 2Å	I-rmsd > 2Å
LISA	**0.83**	**0.84**	**0.78**	**0.9**
ZAPP	0.62	0.58	0.63	0.75
Prodigy	0.54	0.63	0.44	0.55
FIREDOCK_EI	0.46	0.59	0.3	0.6
INSIDE	0.4	0.57	0.23	0.3
ConsBind	0.56	0.64	0.43	0.56
———————	—	—	—	—
AP_DCOMPLEX	0.45	0.5	0.39	0.7
AP_dDFIRE	0.14	0.29	0.08	0.24
AP_DFIRE2	0.17	0.33	0.09	0.27
ZRANK	0.33	0.42	0.26	0.75
ZRANK2	0.4	0.36	0.37	0.71
ROSETTADOCK	0.32	0.32	0.23	0.65
PYDOCK_TOT	0.42	0.32	0.47	0.69
SIPPER	0.21	0.25	0.26	0.21
AP_PISA	0.35	0.27	0.35	0.72
FIREDOCK	0.41	0.56	0.24	0.74
FIREDOCK_AB	0.36	0.47	0.2	0.8
CP_PIE	0.44	0.52	0.35	0.57

Table S6. Correlations organised by I-rmsd ranges characterizing rigid, medium-rigid and flexible complexes in AffinityB2_rel*. Correlations between experimental binding energy and predictions obtained with several tools and organised by different I-rmsd intervals on AffinityB2* complexes. Recall that complexes with I-rmsd I-rmsd > 2Å, 1Å≤ I-rmsd ≤ 2Å and I-rmsd < 1Å are classified as flexible, medium-rigid and rigid.

I-rmsd cutoff	AffinityB2_rel	I-rmsd ≤ 1Å	I-rmsd > 1Å
LISA	**0.81**	**0.8**	**0.82**
ZAPP	0.61	0.64	0.6
Prodigy	0.55	0.65	0.47
FIREDOCK_EI	0.46	0.6	0.37
INSIDE	0.39	0.55	0.26
———————	—	—	—
AP_DCOMPLEX	0.44	0.49	0.44
AP_dDFIRE	0.14	0.3	0.1
AP_DFIRE2	0.17	0.33	0.12
ZRANK	0.34	0.46	0.34
ZRANK2	0.39	0.37	0.45
ROSETTADOCK	0.32	0.34	0.32
PYDOCK_TOT	0.44	0.4	0.52
SIPPER	0.22	0.28	0.14
AP_PISA	0.38	0.37	0.43
FIREDOCK	0.42	0.59	0.35
FIREDOCK_AB	0.37	0.5	0.34
CP_PIE	0.44	0.55	0.39

Table S7. Correlations organised by I-rmsd ranges characterizing rigid and flexible complexes in AffinityB2_rel. Correlations between experimental and predicted binding energies obtained with LISA and 17 other tools and organised by different I-rmsd intervals on AffinityB2_rel complexes. Recall that complexes with I-rmsd ≤ 1Å and I-rmsd > 1Å are classified as rigid and flexible in (Vangone & Bonvin 2015).

I-rmsd cutoff	AffinityB2_rel*	I-rmsd $\leq$ 1Å	I-rmsd $>$ 1Å
LISA	**0.83**	**0.84**	**0.82**
ZAPP	0.62	0.58	0.67
Prodigy	0.54	0.63	0.47
FIREDOCK_EI	0.46	0.59	0.37
INSIDE	0.4	0.57	0.26
ConsBind	0.56	0.64	0.47
————	—	—	—
AP_DCOMPLEX	0.45	0.5	0.45
AP_dDFIRE	0.14	0.29	0.1
AP_DFIRE2	0.17	0.33	0.12
ZRANK	0.33	0.42	0.34
ZRANK2	0.4	0.36	0.45
ROSETTADOCK	0.32	0.32	0.32
PYDOCK_TOT	0.42	0.32	0.53
SIPPER	0.21	0.25	0.15
AP_PISA	0.35	0.27	0.42
FIREDOCK	0.41	0.56	0.35
FIREDOCK_AB	0.36	0.47	0.34
CP_PIE	0.44	0.52	0.39

Table S8. Correlations organised by I-rmsd ranges characterizing rigid and flexible complexes in AffinityB2_rel*. Correlations between experimental and predicted binding energies obtained with LISA and 17 other tools and organised by different I-rmsd intervals on AffinityB2_rel* complexes. Recall that complexes with I-rmsd $\leq$ 1Å and I-rmsd $>$ 1Å are classified as rigid and flexible in (Vangone & Bonvin 2015).

Intersection	Kastritis81	I-rmsd $>$ 1Å	I-rmsd $\leq$ 1Å
LISA	**0.86**	**0.87**	**0.87**
Prodigy	0.73	0.75	0.74

Table S9. Correlations organised by I-rmsd ranges characterizing rigid and flexible complexes in Kastritis81. Correlations between experimental and predicted binding energies obtained with LISA and Prodigy and organised by different I-rmsd intervals on Kastritis81 complexes. Recall that complexes with I-rmsd $\leq$ 1Å and I-rmsd $>$ 1Å are classified as rigid and flexible in (Vangone & Bonvin 2015).

Complex type	AffinityB2_rel	A	AB	EI	ER	ES	OG	OR	OX
LISA	**0.81**	**0.51**	0.63	0.55	**0.99**	0.78	**0.88**	**0.91**	**0.75**
ZAPP	0.61	0.01	0.13	**0.63**	0	0.29	0.83	0.62	0.3
Prodigy	0.55	0.15	0.56	0.47	0.66	**0.92**	0.77	0.47	0.39
FIREDOCK_EI	0.46	0.05	0.11	0.53	0.81	0.46	0.21	0.65	0.34
INSIDE	0.39	0.32	**0.69**	0.52	0.76	0.42	0	0.62	0.03
———	—	—	—	—	—	—	—	—	—
AP_DCOMPLEX	0.44	0.04	0.43	0.21	0.7	0.81	0.62	0.64	0.37
AP_dDFIRE	0.14	0.25	0.26	0.26	0.74	0.52	0.44	0.24	0.36
AP_DFIRE2	0.17	0.25	0.2	0.25	0.78	0.51	0.65	0.28	0.37
ZRANK	0.34	0.31	**0.94**	0.24	0.9	0.53	0.87	0.56	0.31
ZRANK2	0.39	0.14	0.35	0.19	0.74	0.3	0.83	0.65	0.28
ROSETTADOCK	0.32	0.02	0.05	0.19	0.02	0.13	0.59	0.61	0.21
PYDOCK_TOT	0.44	0.07	0.69	0.38	0.82	0.19	0.84	0.69	0.37
SIPPER	0.22	0	0.53	0.02	0.64	0.48	0.35	0.36	0.14
AP_PISA	0.38	0.19	1	0.45	0.61	0.04	0.67	0.56	0.34
FIREDOCK	0.42	0.29	0.81	0.5	0.79	0.45	0.6	0.6	0.38
FIREDOCK_AB	0.37	0.27	0.89	0.37	0.75	0.44	0.71	0.58	0.36
CP_PIE	0.44	0.22	0.8	0.29	0.73	0.49	0.53	0.63	0.47

Table S10. Correlations organised by functional classes. Correlation values computed for several tools and for the eight functional classes in AffinityB2_rel: antibody/antigen (A or AB with bound antibody); enzyme/inhibitors (EI); enzyme/substrate (ES); enzyme/regulatory subunit (ER); G-protein containing (OG); membrane receptors (OR); and others (OX).

Complex type	AffinityB2_rel*	A	AB	EI	ER	ES	OG	OR	OX
LISA	**0.83**	**0.51**	-0.4	**0.69**	**0.99**	0.78	**0.88**	**0.91**	**0.75**
ZAPP	0.62	0.01	0.69	0.58	0.09	0.29	0.83	0.62	0.43
Prodigy	0.54	0.15	0.73	0.43	0.67	**0.92**	0.77	0.47	0.4
FIREDOCK_EI	0.46	0.05	**1**	0.62	0.81	0.46	0.21	0.65	0.34
INSIDE	0.4	0.32	0.88	0.56	0.77	0.42	0	0.62	0.02
ConsBind	0.56	0.3	0.78	0.35	0.84	0.86	0.67	0.61	0.37
———	—	—	—	—	—	—	—	—	—
AP_DCOMPLEX	0.45	0.04	**1**	0.27	0.74	0.81	0.62	0.64	0.37
AP_dDFIRE	0.14	0.25	0.37	0.25	0.74	0.52	0.44	0.24	0.35
AP_DFIRE2	0.17	0.25	0.37	0.25	0.77	0.51	0.65	0.28	0.36
ZRANK	0.33	0.31	0.89	0.22	0.9	0.53	0.87	0.56	0.31
ZRANK2	0.4	0.14	0.44	0.24	0.74	0.3	0.83	0.65	0.27
ROSETTADOCK	0.32	0.02	0.07	0.19	0.02	0.13	0.59	0.61	0.23
PYDOCK_TOT	0.42	0.07	0.15	0.31	0.81	0.19	0.84	0.69	0.36
SIPPER	0.21	0	**1**	0.02	0.9	0.48	0.35	0.36	0.13
AP_PISA	0.35	0.19	0.99	0.35	0.58	0.04	0.67	0.56	0.34
FIREDOCK	0.41	0.29	0.14	0.54	0.78	0.45	0.6	0.6	0.38
FIREDOCK_AB	0.36	0.27	0.31	0.39	0.73	0.44	0.71	0.58	0.37
CP_PIE	0.44	0.22	0.21	0.28	0.75	0.49	0.53	0.63	0.46

Table S11. Correlations organised by functional classes. Correlation values computed for several tools and for the eight functional classes in the intersection between AffinityB2_rel and ConsBind dataset (AffinityB2_rel*): antibody/antigen (A or AB with bound antibody); enzyme/inhibitors (EI); enzyme/substrate (ES); enzyme/regulatory subunit (ER); G-protein containing (OG); membrane receptors (OR); and others (OX).

	One thread	Four threads
LISA	9.07 hrs	3.68 hrs
LISA/NCI	13.4 min	/
Prodigy	3.58 min	/

Table S12. LISA computational time performance. The table shows the performance of LISA compared to Prodigy on the set of 125 complexes in AffinityB2_rel. Computational performance is computed by using a single thread or four threads. Note that NCI-PLOT takes advantage of multithreading computation. Prodigy performance could be compared to LISA because of the available software.

	AffinityB2_rel (%)	Kastritis_rel (%)	Vreven_rel (%)
A	16	11.22	33.33
AB	3.2	4.08	0
EI	17.6	18.37	14.81
ER	6.4	6.12	7.41
ES	6.4	8.16	0
OG	10.4	13.27	0
OR	13.6	12.24	18.52
OX	26.4	26.53	25.93
—	—	—	—
Rigid	52	53.06	48.15
Medium-rigid	36	34.69	40.74
Flexible	12	12.24	11.11

Table S13. Functional and structural composition of AffinityB2_rel, Kastritis_rel and Vreven_rel datasets. Subdivisions in functional classes of AffinityB2_rel, Kastritis_rel and Vreven_rel datasets. Functional classes are: antibody/antigen (A or AB with bound antibody); enzyme/inhibitors (EI); enzyme/substrate (ES); enzyme/regulatory subunit (ER); G-protein containing (OG); membrane receptors (OR); and others (OX). The last three rows in the table show the different composition for AffinityB2_rel, Kastritis_rel and Vreven_rel following the rigid (I-rmsd < 1 Å), medium-rigid (1 Å$\leq$ I-rmsd ≤ 2 Å) and flexible (I-rmsd > 2 Å) classification used in (Vreven et al. 2012).

	AffinityB2_rel	Kastritis_rel	Vreven_rel
A	20	11	9
AB	4	4	0
EI	22	18	4
ER	8	6	2
ES	8	8	0
OG	13	13	0
OR	17	12	5
OX	33	26	7
—	—	—	—
Rigid	65	52	13
Medium-rigid	45	34	11
Flexible	15	12	3

Table S14. Number of complexes organised in functional and structural classes. Complexes in the AffinityB2_rel, Kastritis_rel and Vreven_rel datasets are counted with respect to their function and their flexibility. For the list of functions, see legend of **Table S13**.