

Evaluating the Difficulty Involved in Designing Small Molecule Drugs to Inhibit Protein-Protein Interactions

Delaram Ahmadi and David Barlow*

Department of Pharmacy, King's College London, London SE1 9NH, UK

Article history:

Received: 17 July 2017

Accepted: 24 July 2017

HIGHLIGHTS

- Mean interfacial surface curvatures have been determined for protein-protein interaction (PPI) partners in their complexed and uncomplexed states.
- Mean interfacial surface roughnesses have been determined for protein-protein interaction (PPI) partners in their complexed and uncomplexed states.
- Amino acid compositions have been determined for PPI interface surfaces and these compared with that for the average protein surface.
- Quantification of the PPI interfacial surface properties is used to assess the druggability of these targets.

ABSTRACT

Keywords:

Drug Design
Druggability
Molecular Modelling
Protein-protein interactions

The targeting of drugs to block protein-protein interactions (PPIs) has attracted great interest over recent years. Such targets, however, have been held to be difficult to inhibit using low molecular weight compounds, and as a consequence they are often branded as “undruggable”. This is partly because the interfaces involved are seen to be large, and the fact that they are generally regarded as being too smooth and too flat. In the work reported here, a series of quantitative systematic studies have been performed to determine the molecular area, roughness, curvature, and amino acid composition of the interfacial surfaces of PPIs, to determine the feasibility of designing small molecule drugs to inhibit these interactions. The X-ray crystal structures are analysed for a set of 48 PPIs involving G-protein, membrane receptor extracellular domain, and enzyme-inhibitor complexes. The protein partners involved in these PPIs are shown to have much larger interfacial areas than those for protein-small molecule complexes ($\geq 900 \text{ \AA}^2$ vs $\sim 250 \text{ \AA}^2$ respectively), and they have interfaces that are fairly smooth (with fractal dimensions close to 2) and quite flat (with mean surface curvatures in the order of $\pm 0.1 \text{ \AA}^{-1}$). The mean interfacial surface curvatures of the PPI protein partners, however, are seen to change upon complexation, some very significantly so. Despite the fact that the amino acid compositions of the PPI interface surfaces are found to be significantly different from that of the average protein surface (with variations according to the type of PPI), it is concluded that the prospects for designing low molecular weight PPI inhibitors that act in an orthosteric manner remain rather limited.

* Corresponding Author:

Email: dave.barlow@kcl.ac.uk (D. Barlow)

Introduction

Protein-protein interactions (PPIs) play a fundamental role in mediating key biological processes, such as the mammalian immune response, cell-cell recognition, cell proliferation, growth, differentiation, signal transduction and apoptosis. Changes in the specificity and affinity of these interactions can lead to cellular malfunction like uncontrolled cell growth that results in cancer (Toogood, 2002; White et al., 2008; Raj et al., 2012; Nero et al., 2014). Inevitably, interest has arisen amongst scientists within the pharmaceutical industry to develop inhibitors of these interactions as therapeutic agents.

In the past decade, therapeutic proteins such as monoclonal antibodies, which target PPIs that are accessible to the extracellular environment, have proven tremendously successful. However, these macromolecules can be expensive and are rarely orally bioavailable. Thus, the relative lower cost, ease of use and potential to target intracellular PPIs all favour the discovery of small synthetic molecule (<500 Da) PPI modulators (Nero et al., 2014).

However, PPIs occur over a large surface area, with approximately 750 to 1,500 Å² buried on each side of the protein interface (Jones and Thornton, 1996; Conte et al., 1999). Moreover, the contact surfaces involved at these interfaces are qualitatively observed to be rather smooth, flat and featureless, and thus lack the deeply buried grooves and pockets that provide for the high affinity binding sites for small molecules like those found in the substrate binding sites of enzymes (Wells and McClendon, 2007).

Despite the ‘undruggability’ of these PPIs, recent studies show that there have been some successes seen in the development of small molecule inhibitors as cancer therapeutics. For example, with the recent case of the MDM2/p53 interaction inhibitors, there are interaction ‘hot spots’ on the PPI surfaces detected, which can provide a focus for small molecule inhibitor interactions [Conte et al., 1999; Vassilev et al., 2004; Morrow and Zhang, 2012; Thangudu et al., 2012]. At the same time, there has been much ongoing research in the challenges governing the inhibition of PPIs with small molecule drugs (Arkin and Wells, 2004; Laraia et al., 2015).

The complexes of interest within this study – specifically, those reflecting PPIs which might be targeted in cancer therapy – were chosen from those identified by Kastiris et al., (2011) and are binary complexes involving either G-proteins (OG), receptor extracellular domains (OR) or enzyme inhibitors (EI).

The research reported here focuses on providing a systematic quantitative analysis of the characteristics of these PPI surfaces and the structural changes involved at the interfaces following complexation. By testing the hypothesis that the PPIs are relatively extensive in interface area, have smooth interfaces that are relatively

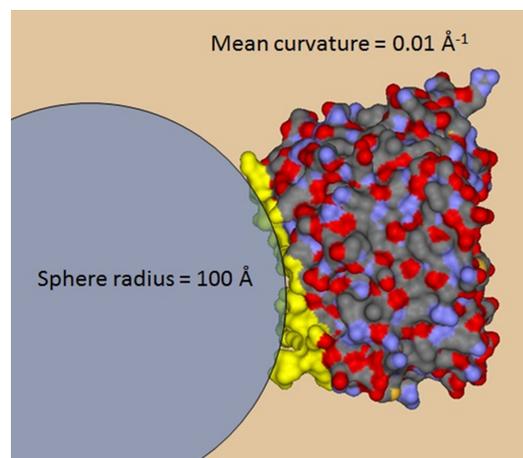


Figure 1. Cartoon illustrating the relation between the mean surface curvature values (k) and the notional external sphere fitted to the surface, with radius = k^{-1} . Negative values of k indicate convex surfaces, positive values indicate concave surfaces.

flat, we aimed to evaluate the difficulty in devising low molecular weight drugs targeted against PPIs of the kind relevant to anticancer drug discovery.

Materials and Methods

A database of 48 PPIs (with X-ray single crystal structures determined to ≤ 2.5 Å resolution) was compiled from the information provided by Kastiris et al., (2011). The atomic co-ordinates of the protein complexes of interest (together with those of the unbound protein partners) were obtained from the Protein Data Bank in Europe (PDBe) (Gutmanas et al., 2013) accessed at <http://www.ebi.ac.uk/pdbe/>. The molecular surface areas and amino acid compositions of the PPI interfaces for the bound and unbound proteins involved in G-protein (OG) complexes, receptor extracellular domain (OR) complexes, and enzyme inhibitor (EI) complexes were obtained using PDBePISA (Krissinel et al., 2007). Mean interfacial surface curvature values for the component proteins involved in PPI complexes were computed (in Å⁻¹, see Fig. 1) using the SurfRace program (Tsodikov et al., 2002), employing a probe radius of 1.2 Å.

The roughness of each PPI interface was computed from the fractal dimension of the surface (D), using the gradient ($2 - D$) of a plot of $\log(A)$ vs. $\log(r)$, (Lewis and Rees, 1985) where A is the molecular surface area of the interface atoms (calculated using SurfRace) given a spherical probe of radius, r (varied over the range 0.75 Å to 4.0 Å), using the atomic van der Waals radii given by Chothia (1976). D can take values from 2 (which indicates a perfectly smooth surface) through to 3 (which indicates

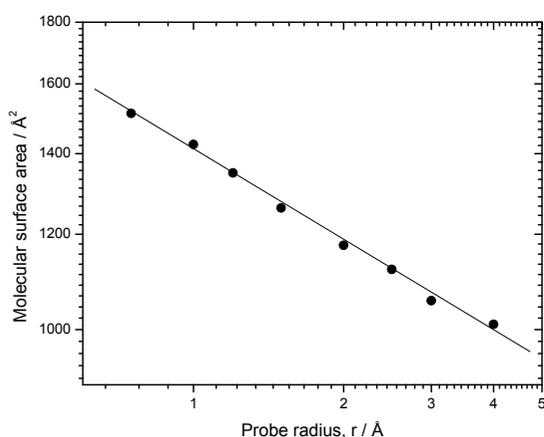


Figure 2. PPI interfacial molecular surface area (\AA^2) of the uncomplexed form of ADP-ribosylation factor 1 (PDB code, 1J2J) as a function of probe radius, r (\AA). The slope of the plot allows determination of the surface fractal dimension, D , in this case 2.25.

a very rough surface).

Mean interfacial surface curvature and roughness values were only calculated for PPI protein partners whose structures in the unbound state were determined by X-ray crystallography. In cases where the structure of the unbound form was determined by NMR, the pairwise comparisons between its bound and unbound species were not performed.

The generic mean frequencies of amino acids across protein surfaces were taken from Miller et al., (1987), and the frequencies of the residue types found within PPI interfaces were computed using the PDBGoodies utility (Hussain et al., 2002).

Pairwise statistical analyses were performed using a Student's *t*-test, and three-way comparisons using 1-way ANOVA, with statistical significance taken as $p < 0.05$.

Results and Discussion

Tables 1 and 2 present the details of the 48 protein-protein complexes considered for analysis: Table 1 shows details for the 12 G-protein (OG) and 7 receptor extracellular domain complexes (OR) complexes, and Table 2 shows the details for the 29 enzyme inhibitor (EI) complexes.

The interface areas of the PPIs (calculated as half the difference in the surface areas of the complexed and uncomplexed protein partners), ranged from 476 \AA^2 to 1712 \AA^2 , and were significantly greater, therefore, than those for protein-small molecule interactions, which average around 500 \AA^2 (Wells & McClendon, 2007). There were modest but statistically significant differences found for the interfacial areas of the OG and OR complexes versus the EI complexes, with the mean areas determined as $952 (\pm 415) \text{ \AA}^2$ and $875 (\pm 229) \text{ \AA}^2$, respectively.

Log-log plots of the variation in PPI interface surface area as a function of probe radius were in all cases linear (with $r^2 \geq 0.95$; *cf.*, Fig. 2), and allowed reliable determination of the interfacial surface irregularity/surface roughness, described by means of the associated fractal dimension (D). All the PPI interface surfaces were shown to be smooth (Table 3; Fig. 3), with the mean fractal dimensions for their unbound protein components of 2.25, 2.18 and 2.23 for OG, OR and EI complexes, respectively. One-way ANOVA analysis for the variation in the fractal dimensions of the unbound components of the OG, OR and EI complexes showed no significant difference between the mean interfacial surface roughnesses of the three types of complex.

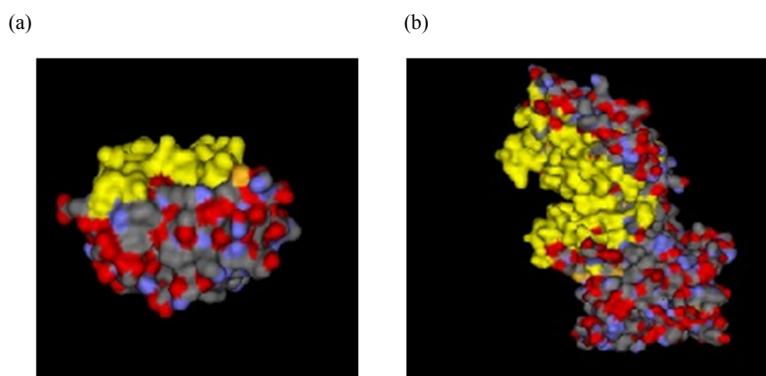


Figure 3. Space-filling models showing the ligand binding domain of the ephrin type B receptor 4 (a; PDB code, 2HLE) and the catalytic domain of P120 gap protein (b; PDB code 1WQ1), with the PPI interface surfaces highlighted yellow. The interface surface of the ephrin domain is fairly smooth, with a fractal dimension of $D = 2.25$, and the P120 protein interface surface is somewhat more rough, with $D = 2.46$.

Table 1. Data set of G-protein (OG) and receptor extracellular domain (OR) complexes with their unbound components and corresponding structural and interfacial properties.

Class	Complex (PDB entry)	Interface Area (Å ²)	Resolution (Å) ^b	Protein partner 1		Resolution (Å) ^b	Number of residues	Protein partner 2		Resolution (Å) ^b	Number of residues
				PDB entry ^a	Name			PDB entry ^a	Name		
OG	1A2K	1595	2.50	1QG4 (D)	GTP-binding nuclear protein Ran	2.50	22	1OUN (A)	Nuclear transport factor 2	2.30	27
OG	1E96	590	2.40	1MH1 (A)	Ras-related C3 botulinum toxin substrate	1.38	18	1HH8 (B)	Neutrophil cytosol factor 2	1.80	20
OG	1FQJ	899	2.02	1TND (A)	Guanine nucleotide-binding protein G(t) subunit alpha-1	2.20	24	1FQI (B)	Regulator of G-protein signalling 9	1.94	28
OG	1I2M	1389	1.76	1QG4 (A)	GTP-binding nuclear protein Ran	2.50	34	1A12 (B)	Regulator of chromosome condensation	1.70	47
OG	1I4D	703	2.50	1I49 (A)	Arfaptin-2	2.80	19	1MH1 (D)	Ras-related C3 botulinum toxin substrate 1	1.38	20
OG	1IBR	1695	2.30	1QG4 (C)	GTP-binding nuclear protein Ran	2.50	49	1F59 (D)	Importin subunit beta-1	2.80	54
OG	1J2J	604	1.60	1O3Y (A)	ADP-ribosylation factor 1	1.50	18	1OXZ (B)	ADP-ribosylation factor-binding protein GGA1	1.50	14
OG	1LFD	619	2.10	5P21 (D)	GTPase HRas	1.35	18	1LXD (C)	Ral guanine nucleotide dissociation stimulator	2.40	15
OG	1NVU	1712	2.20	1LF0 (R)	GTPase HRas	1.70	43	2II0 (S)	Son of sevenless homolog 1 variant (Fragment)	2.02	57
OG	1WQ1	1420	2.50	6Q21 (R)	GTPase HRas	1.95	37	1WER (G)	Ras GTPase-activating protein 1	1.60	45
OG	1Z0K	927	1.92	2BME (C)	Ras-related protein Rab-4A	1.57	27	1YZM (D)	Rabenosyn-5	1.50	24
OG	2FJU	626	2.20	2ZKM (B)	Phospholipase C-beta-2	1.62	19	1MH1 (A)	Ras-related C3 botulinum toxin substrate 1	1.38	14
OR	1E4K	477	2.30	2DTQ (B)	Ig gamma-1 chain C region	2.00	15	1FNL (C)	Fc-gamma RIIIb	1.80	12
OR	1EER	975	1.90	1BUY (A)	Erythropoietin	NMR	25	1ERN (B)	Erythropoietin receptor	2.40	28
OR	1KTZ	493	2.15	1TGK (A)	Transforming growth factor beta-3	3.30	10	1M9Z (B)	TGF-beta receptor type-2	1.05	15

Table 1. Continued.

Class	Complex (PDB entry)	Interface Area (Å ²)	Resolution (Å) ^b	Protein partner 1		Resolution (Å) ^b	Number of residues	Protein partner 2		Resolution (Å) ^b	Number of residues
				PDB entry ^a	Name			PDB entry ^a	Name		
OR	1T6B	967	2.50	1ACC (X)	Protective antigen	2.10	32	1SHU (Y)	Anthrax toxin receptor 2	1.50	26
OR	1XU1	646	1.90	1U5Y (B)	TNF domain of APRIL	0.96	32	1XUT (S)	TNFR 13B TACI CRD2 domain	NMR	14
OR	2HLE	1052	2.05	2BBA (A)	Ephrin type-B receptor 4	1.65	37	1IKO (B)	Ephrin-B2	1.92	25

a) PDB entry codes of the protein-protein complexes or their unbound protein partners, with the chain identifiers for partners 1 and 2 given in parentheses.

b) Where the 3-D structure of an unbound protein partner has been determined by NMR, this is indicated; otherwise, the resolution of the X-ray structure determination is given.

A comparison of the mean interfacial surface roughness of the bound and uncomplexed protein partners involved in OR and EI complexes showed no statistically significant difference, but for the OG complexes, the interfacing surfaces of the partner proteins were found to become a little less rough on complexation, with the mean interfacial surface fractal dimension decreasing from $D = 2.25$ to $D = 2.21$ ($p = 0.015$).

All the PPI interfaces studied were likewise determined to be relatively flat, with their mean surface curvatures lying in the range, $-0.12 \text{ \AA}^{-1} - 0.11 \text{ \AA}^{-1}$ (Table 4). For most of the proteins involved in these PPIs (92%), however, there are differences between the mean interfacial surface curvatures of their complexed and uncomplexed forms, some of which – like phospholipase C3 (2FJU), staphopain B (1PXV), and Ras-related Rab-

4A (1Z0K) – are quite substantial. The mean absolute difference in surface curvatures for all the complexed and uncomplexed species is 0.028.

Statistical comparisons of the frequencies of the 20 different amino acid types across protein surfaces in general and their frequencies within PPI interfaces revealed significant differences between OG, OR and EI complexes, and the findings for the former two classes of complex are shown graphically in Fig. 4. Arginine and leucine show a statistically significant enrichment at the PPI interfaces of OG and OR complexes, whereas EI complexes show a significant enrichment of histidine, leucine and cysteine (data not shown). The residue types that are significantly underrepresented at PPI interfaces tend to be those with small and/or polar side chains, including alanine and glycine in the PPI interfaces of OR

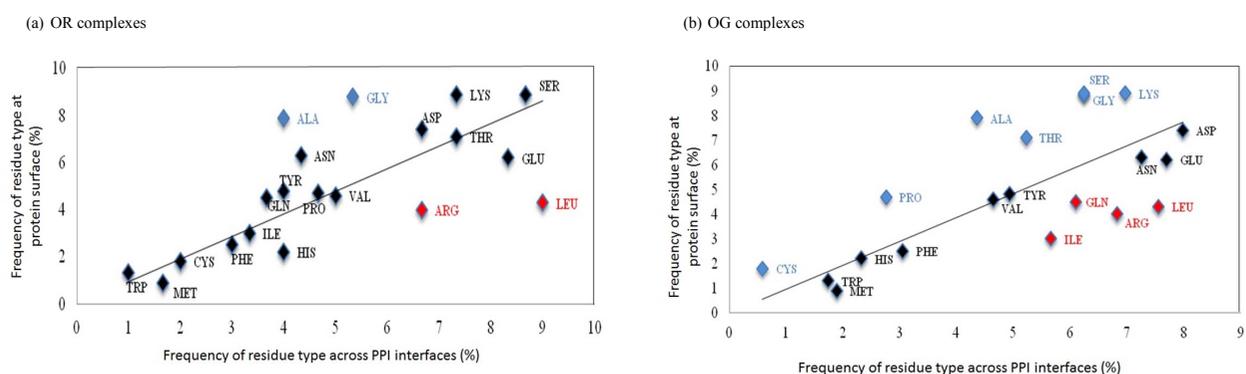


Figure 4. Percentage frequencies of amino acid residue types at protein surfaces and across PPI interfaces. Residue types showing a statistically significant enrichment across PPI interfaces are highlighted red; those showing statistically significant depletion at PPI interfaces are highlighted blue. Data relate to OR complexes (a) and OG complexes (b).

Table 2. Data set of enzyme inhibitor (EI) complexes with their unbound components and corresponding structural and interfacial properties.

Class	Complex (PDB entry)	Protein partner 1			Protein partner 2			Number of residues		
		Interface Area (\AA^2)	Resolution (\AA) ^b	PDB entry ^a	Name	Resolution (\AA) ^b	PDB entry ^a		Name	
EI	1ACB	770.8	2.00	4CHA (E)	Chymotrypsinogen A	1.68	1EGL (I)	Eglin C	NMR	17
EI	1AVX	792.1	1.90	1QQU (A)	Trypsin	1.63	1BA7 (B)	Trypsin inhibitor A	2.50	20
EI	1BRS	778.5	2.00	1A2P (A)	Ribonuclease	1.50	1A19 (D)	Barstar	2.76	20
EI	1BVN	1106.6	2.50	1PIG (P)	Pancreatic alpha-amylase	2.20	1HOE (T)	Alpha-amylase inhibitor	2.00	28
EI	1EMV	766.8	1.70	1FSJ (B)	Colicin-E9	1.80	1IMQ (A)	Colicin-E9 immunity protein	NMR	81
EI	1EZU	934.0	2.40	1TRM (D)	Anionic trypsin-2	2.30	1ECZ (A)	Ecotin	2.68	30
EI	1F34	1567.1	2.45	4PEP (A)	Pepsin A	1.80	1F32 (B)	Major pepsin inhibitor 3	1.75	45
EI	1FLE	883.5	1.90	9EST (E)	Chymotrypsin-like elastase family member 1	1.90	2REL (I)	Elafin	NMR	20
EI	1GL1	799.1	2.1	4CHA (A)	Chymotrypsinogen A	1.68	1PMC (I)	Protease inhibitor	NMR	20
EI	1HIA	631.2	2.40	2PKA (B)	Glandular kallikrein	2.05	1BX8 (I)	Hirustasin	1.40	14
EI	1JIW	1004.1	1.74	1AKL (P)	Alkaline metalloproteinase	2.00	2RN4 (I)	Proteinase inhibitor	NMR	24
EI	1JTG	1297.7	1.73	3GMU (B)	Beta-lactamase inhibitory protein	1.98	1ZG4 (A)	Beta-lactamase TEM	1.55	35
EI	1NB5	894.1	2.40	8PCH (D)	Pro-cathepsin H	2.10	1DVC (L)	Cystatin-A	NMR	31
EI	1OPH	677.0	2.30	1QLP (A)	Alpha-1-antitrypsin	2.0	2PTN (B)	Cationic trypsin	1.55	31
EI	1PPE	842.2	2.00	2PTN (E)	Cationic trypsin	1.55	1LU0 (I)	Trypsin inhibitor 1	1.03	18
EI	1PXV	1164.1	1.80	1X9Y (A)	Staphopain B	2.50	1NYC (C)	Staphostatin B	1.40	30

Table 2. Continued.

Class	Complex (PDB entry)	Interface Area (\AA^2)	Protein partner 1			Protein partner 2			Number of residues
			Resolution (\AA) ^b	PDB entry ^a	Name	Resolution (\AA) ^b	PDB entry ^a	Name	
EI	1R0R	705.3	1.10	1SCN (E)	Subtilisin Carlsberg	1.90	2GKR (I)	Ovomucoid	15
EI	1UUG	1056.7	2.40	3EUG (C)	Uracil-DNA glycosylase	1.43	2UGI (D)	Uracil-DNA glycosylase inhibitor	32
EI	2ABZ	712.6	2.16	3IUU (A)	Carboxypeptidase A1	1.39	1ZFI (C)	Metallo-carboxypeptidase inhibitor	19
EI	2B42	1259.9	2.50	2DCY (B)	Xylanase	1.40	1T6E (A)	Xylanase inhibitor	41
EI	2J0T	774.8	2.54	966C (B)	Interstitial collagenase	1.90	1D2B (E)	Metalloproteinase inhibitor 1	20
EI	2O3B	815.7	2.30	1ZM8 (A)	Nuclease	1.90	1J57 (B)	Sugar-non-specific nuclease inhibitor	19
EI	2OUL	966.9	2.20	3BPF (A)	Falcpain 2	2.90	2NNR (B)	Chagasin	28
EI	2PTC	714.3	1.90	2PTN (E)	Cationic trypsin	1.55	9PTI (I)	Pancreatic trypsin inhibitor	15
EI	2SIC	806.9	1.80	1SUP (E)	Subtilisin BPN'	1.60	3SSI (I)	Subtilisin inhibitor	15
EI	2SNI	813.9	2.10	1UBN (E)	Subtilisin BPN'	2.40	2CI2 (I)	Subtilisin-chymotrypsin inhibitor-2A	18
EI	2UUY	637.2	1.15	2PTN (A)	Cationic trypsin	1.55	2UUX (B)	Trypsin inhibitor	12
EI	3SGB	633.2	1.80	2QA9 (E)	Streptogrisin-B	1.18	2OYO (I)	Ovomucoid	14
EI	4CPA	558.7	2.50	8CPA (B)	Carboxypeptidase A1	2.00	1H20 (J)	Metallo-carboxypeptidase inhibitor	12

a) PDB entry codes of the protein-protein complexes or their unbound protein partners, with the chain identifiers for partners 1 and 2 given in parentheses.

b) Where the 3-D structure of an unbound protein partner has been determined by NMR, this is indicated; otherwise, the resolution of the X-ray structure determination is given.

complexes, alanine and aspartic acid in EI complexes (data not shown), and alanine, glycine, and serine in OG complexes.

Previous systematic analyses of PPIs have sometimes included consideration of complexes whose crystal structures were determined to resolutions $> 2.5 \text{ \AA}$ (Jones and Thornton, 1996; Conte et al., 1999), and the dataset selected for study here was deliberately restricted to those whose structures were refined at medium to high resolution, in order to give greater confidence in atomic co-ordinates and any quantities derived from these. Many of the earlier analyses of PPIs also considered a wider spectrum of interaction types than was considered here – including those between the sub-units of oligomeric proteins, and those between antigens and antibodies (Jones and Thornton, 1996; Conte et al., 1999). Given that these PPIs interfaces are of limited relevance as regards the design of (intracellularly acting) anticancer drugs, they too were excluded from consideration in the studies reported here.

Notwithstanding these restrictions to the dataset, however, and in full accord with the observations recorded by previous workers, we find that the interfaces of the PPIs included in our dataset cover areas around 5 to 10 times larger than those seen in complexes involving proteins and low molecular ligands. Moreover, these interfaces are seen to have surfaces that are quite smooth and flat, with very little change in the roughness of the relevant surfaces of the protein partners in their transition from their unbound to bound states.

Targeting of these interfaces using non-protein ligands that bind across the whole interface area would thus require compounds of relatively high molecular weight, which would then preclude their administration *via* the oral route (given that such drugs would ideally have molecular weights $< 500 \text{ Da}$ (Lipinski et al., 2001)). In addition, because the PPI interfaces are so smooth (and also quite flat), these high molecular weight ligands would of necessity bind to the target protein surface with around half of their structure remaining exposed to the surrounding solvent, and this would greatly limit the entropic contribution to the free energy of their complexation provided through desolvation. While such a deficiency might be mitigated if the PPI interfaces involved significant numbers of charged and polar groups which would allow for extensive electrostatic and hydrogen bond interactions, our analysis of the amino acid composition of the PPI interface surfaces shows that this is not generally the case. Most of the PPI interface surfaces are depleted in small polar residues (like serine, threonine), and they are also often deficient in the charged residues, lysine and glutamic acid, although the interface surfaces in OG and OR complexes are enriched in arginine, and EI complexes are enriched in histidine.

It has been proposed, however, that low molecular

weight PPI inhibitors of high affinity can be successfully achieved provided they are targeted to bind to the interfaces' key 'hot spot' residues (Bogan and Thorn, 1998; Wells and McClendon, 2007; Thangudu et al., 2012; Cukuroglu et al., 2014). It should be noted though that even if this is so, the resulting enthalpic contribution to the free energy of complexation is very unlikely to compensate the fact that van der Waals interactions will be far fewer than for similar-sized ligands binding within buried pockets, and unlikely also to compensate for the minimal contribution made to binding affinity by ligand desolvation.

To further compound the issue, we find that virtually all the proteins involved in the PPIs studied here undergo changes in their mean interfacial surface curvature upon complexation, and these changes – taken together with the changes in interface residue conformation shown by Kastiris et al. (2011) – would pose problems not only for the identification of so-called 'hot spot' residues but problems too for the quantitative assessment of putative ligand binding affinities.

Conclusions

We thus note that while various reviews have reported recent good progress in the design of PPI inhibitors (Corbi-Verge and Kim, 2016), with successes cited where molecules have been taken into clinical trials (Nero et al., 2014; Laraia et al., 2015), we note too that the molecular weights of these compounds are still quite high (typically around 1000 to 1500 Da), and all of the compounds (assuming they are even successful in trials) are thus likely to be formulated for parenteral (and not oral) use. The large size of these compounds can also be expected to result in them having a rather low membrane permeability, and their high numbers of hydrogen bond donor and acceptor groups may mean that they will likely satisfy the pharmacophores of a great number of off-target proteins, leading to innumerable side-effects. Moreover, many of these compounds are shown to bind as allosteric (rather than orthosteric) inhibitors, and so their inhibitory activity is explained not by their binding at the PPI interface but by their binding elsewhere on the protein surface and thereby triggering a conformational change in the protein, which then indirectly interferes with the protein-protein interaction. Taken in conjunction with the observations that are recorded here, therefore, the prospects for successful design of orthosteric PPI inhibitors for use in cancer therapy would seem very limited.

Acknowledgements

The authors gratefully acknowledge helpful discussions with Dr Oleg Tsodikov (University of Kentucky, USA), in respect of his SurfRace program.

Table 3. Interfacial surface roughness/fractal dimension (D) of PPI protein partners.

Class	PDB code (chain id)	Component	Fractal Dimension (D)	
			Unbound	Bound
EI	1ACB (E)	Chymotrypsinogen A	2.27	2.41
EI	1AVX (A)	Trypsin	2.34	2.36
EI	1AVX (B)	Trypsin inhibitor A	2.17	2.16
EI	1BRS (A)	Ribonuclease	2.21	2.22
EI	1BVN (P)	Pancreatic alpha-amylase	2.39	2.30
EI	1EMV (B)	Colicin-E9	2.14	2.15
EI	1EZU (D)	Anionic trypsin-2	2.31	2.27
EI	1F34 (A)	Pepsin A	2.18	2.33
EI	1F34 (B)	Major pepsin inhibitor 3	2.23	2.19
EI	1GL1 (A)	Chymotrypsinogen A	2.23	2.31
EI	1HIA (B)	Glandular Kallikrein	2.37	2.41
EI	1HIA (I)	Hirustasin	2.18	2.17
EI	1JIW (P)	Alkaline metalloproteinase	2.33	2.37
EI	1JTG (A)	Beta-lactamase TEM	2.15	2.19
EI	1JTG (B)	Beta-lactamase inhibitory protein	2.19	2.27
EI	1NB5 (D)	Pro-cathepsin H	2.16	2.17
EI	1OPH (A)	Alpha-1-antitrypsin	2.13	2.17
EI	1OPH (B)	Cationic trypsin	2.24	2.35
EI	1PPE (E)	Cationic trypsin	2.31	2.37
EI	1PPE (I)	Trypsin inhibitor 1	2.19	2.14
EI	1PXV (A)	Staphopain B	2.50	2.18
EI	1PXV (C)	Staphostatin B	2.15	2.16
EI	1R0R (E)	Subtilisin Carlsberg	2.22	2.22
EI	1R0R (I)	Ovomucoid	2.13	2.13
EI	1UUG (C)	Uracil-DNA glycosylase	2.17	2.23
EI	1UUG (D)	Uracil-DNA glycosylate inhibitor	2.20	2.25
EI	2ABZ (A)	Carboxypeptidase A1	2.29	2.30
EI	2ABZ (C)	Metalloproteinase inhibitor	2.21	2.19
EI	2B42 (A)	Xylanase inhibitor	2.22	2.30
EI	2B42 (B)	Xylanase	2.23	2.45
EI	2J0T (B)	Interstitial collagenase	2.24	2.25
EI	2O3B (A)	Nuclease	2.19	2.16
EI	2OUL (B)	Chagasin	2.19	2.18

Table 3. Continued.

Class	PDB code (chain id)	Component	Fractal Dimension (D)	
			Unbound	Bound
EI	2PTC (I)	Pancreatic trypsin 33 inhibitor	2.11	2.12
EI	2SIC (E)	Subtilisin BPN'	2.27	2.32
EI	2SIC (I)	Subtilisin inhibitor	2.17	2.09
EI	2SNI (E)	Subtilisin BPN'	2.37	2.30
EI	2SNI (I)	Subtilisin-chymotrypsin inhibitor 2A	2.13	2.13
EI	2UUY (A)	Cationic trypsin	2.23	2.33
EI	2UUY (B)	Trypsin inhibitor	2.08	2.10
EI	3SGB (E)	Streptogrisin-B	2.24	2.24
EI	3SGB (I)	Ovomucoid	2.12	2.09
EI	4CPA (B)	Carboxypeptidase	2.36	2.38
OG	1A2K (A)	Nuclear transport factor 2	2.34	2.32
OG	1A2K (D)	GTP-binding nuclear protein Ran	2.28	2.20
OG	1E96 (A)	Ras-related C3 botulinum toxin	2.14	2.17
OG	1E96 (B)	Neutrophil cytosol factor 2	2.26	2.19
OG	1FQJ (A)	Guanine nucleotide-binding protein G(t) subunit alpha-1	2.22	2.16
OG	1FQJ (B)	Regulator of G-protein signalling 9	2.17	2.24
OG	1I2M (A)	GTP-binding nuclear protein Ran	2.23	2.26
OG	1I2M (B)	Regulator of chromosome condensation	2.29	2.22
OG	1I4D (D)	Ras-related C3 botulinum toxin substrate 1	2.29	2.14
OG	1IBR (C)	GTP-binding nuclear protein Ran	2.28	2.13
OG	1J2J (A)	ADP-ribosylation factor 1	2.21	2.26
OG	1J2J (B)	ADP-ribosylation factor-binding protein GGA1	2.15	2.20
OG	1LFD (C)	Ral guanine nucleotide dissociation stimulator	2.15	2.12
OG	1LFD (D)	GTPase HRas	2.27	2.28
OG	1NVU (S)	Son of sevenless homolog 1 variant (fragment)	2.30	2.34
OG	1NVU (R)	GTPase HRas	2.30	2.19

Table 3. Continued.

Class	PDB code (chain id)	Component	Fractal Dimension (D)	
			Unbound	Bound
OG	1WQ1 (G)	Ras GTPase- activating protein 1	2.46	2.32
OG	1Z0K (C)	Ras-related protein Rab-4A	2.24	2.15
OG	1Z0K (D)	Rabenosyn-5	2.19	2.10
OG	2FJU (A)	Ras-related C3 botulinum toxin substrate 1	2.29	2.23
OG	2FJU (B)	Phospholipase C- beta- 2	2.15	2.16
OR	1E4K (B)	Ig gamma-1 chain C rgeion	2.20	2.16
OR	1E4K (C)	Fc- gamma RIIIb	2.17	2.15
OR	1EER (B)	Erythropoietin receptor	2.23	2.15
OR	1KTZ (B)	TGF- beta receptor type-2	2.11	2.15
OR	1PVH (A)	Interleukin-6 receptor subunit beta	2.22	2.22
OR	1T6B (X)	Protective antigen	2.18	2.23
OR	1T6B (Y)	Anthrax toxin receptor 2	2.17	2.17
OR	1XU1 (B)	TNF domain of April	2.21	2.17
OR	2HLE (A)	Ephrin type-B receptor 4	2.20	2.25
OR	2HLE (B)	Ephrin-B2	2.17	2.21

Table 4. Mean interfacial surface curvatures (\AA^{-1}) of PPI proteins.

Class	PDB entry (chain id)	Complex	Curvature (\AA^{-1})	
			Unbound	Bound
EI	1AVX (A)	Trypsin	0.04	0.07
EI	1AVX (B)	Trypsin inhibitor A	-0.02	-0.02
EI	1F34 (A)	Pepsin A	0.03	0.03
EI	1F34 (B)	Major pepsin inhibitor 3	-0.05	-0.03
EI	1HIA (B)	Glandular Kallikrein	0.11	0.08
EI	1HIA (I)	Hirustasin	-0.05	-0.08
EI	1JTG (A)	Beta-lactamase TEM	0.03	0.07
EI	1JTG (B)	Beta-lactamase inhibitory protein	0.00	0.01
EI	1OPH (A)	Alpha-1-antitrypsin	-0.09	-0.10
EI	1OPH (B)	Cationic trypsin	0.06	0.05
EI	1PPE (E)	Cationic trypsin	0.05	0.06
EI	1PPE (I)	Trypsin inhibitor 1	-0.07	-0.06

Table 4. Continued.

Class	PDB entry (chain id)	Complex	Curvature (\AA^{-1})	
			Unbound	Bound
EI	1PXV (C)	Staphostatin B	-0.09	-0.05
EI	1R0R (E)	Subtilisin Carlsberg	0.03	0.01
EI	1UUG (C)	Uracil-DNA glycosylase	-0.03	-0.06
EI	1UUG (D)	Uracil-DNA glycosylate inhibitor	0.02	0.00
EI	2ABZ (A)	Carboxypeptidase A1	0.05	0.06
EI	2ABZ (C)	Metalloprotease inhibitor	-0.02	-0.04
EI	2B42 (A)	Xylanase inhibitor	0.01	0.01
EI	2B42 (B)	Xylanase	0.00	0.01
EI	2PTC (E)	Cationic trypsin	0.04	0.07
EI	2PTC (I)	Pancreatic trypsin inhibitor	-0.06	-0.05
EI	2SIC (E)	Subtilisin BPN'	0.04	0.02
EI	2SIC (I)	Subtilisin inhibitor	-0.07	-0.09
EI	2SNI (E)	Subtilisin BPN'	0.07	0.04
EI	2SNI (I)	Subtilisin-chymotrypsin inhibitor 2A	-0.07	-0.06
EI	2UUY (A)	Cationic trypsin	0.09	0.06
EI	2UUY (B)	Tryptase inhibitor	-0.12	-0.10
EI	3SGB (E)	Streptogrisin-B	0.02	-0.02
EI	3SGB (I)	Ovomucoid	-0.11	-0.08
OG	1A2K (A)	Nuclear transport factor 2	0.06	0.02
OG	1A2K (D)	GTP-binding nuclear protein Ran	0.01	0.02
OG	1E96 (A)	Ras-related C3 botulinum toxin	-0.06	-0.04
OG	1E96 (B)	Neutrophil cytosol factor 2	0.06	0.04
OG	1FQJ (A)	Guanine nucleotide-binding protein G(t) subunit alpha-1	-0.03	-0.02
OG	1FQJ (B)	Regulator of G-protein signalling 9	-0.05	-0.06
OG	1I2M (A)	GTP-binding nuclear protein Ran	0.01	-0.01
OG	1I2M (B)	Regulator of chromosome condensation	0.04	-0.04
OG	1I4D (D)	Ras-related C3 botulinum toxin substrate 1	-0.09	-0.04
OG	1J2J (A)	ADP-ribosylation factor 1	-0.01	0.04

Table 4. Continued.

Class	PDB entry (chain id)	Complex	Curvature (\AA^{-1})	
			Unbound	Bound
OG	1LFD (C)	Ral guanine nucleotide dissociation stimulator	-0.08	-0.05
OG	1LFD (D)	GTPase HRas	0.04	0.00
OG	1NVU (S)	Son of sevenless homolog 1 variant (fragment)	0.03	0.04
OG	1NVU (R)	GTPase HRas	0.03	-0.01
OG	1WQ1 (R)	GTPase HRas	-0.03	-0.04
OG	1WQ1 (G)	Ras GTPase- activating protein 1	0.02	0.05
OG	1Z0K (D)	Rabenosyn-5	-0.06	-0.07
OG	2FJU (A)	Ras-related C3 botulinum toxin substrate 1	-0.22	-0.06
OG	2FJU (B)	Phospholipase C- beta- 2	-0.05	0.08
OR	1E4K (B)	Ig gamma-1 chain C region	-0.05	-0.06
OR	1E4K (C)	Fc- gamma RIIIb	-0.03	-0.03
OR	1T6B (X)	Protective antigen	-0.03	0.00
OR	1T6B (Y)	Anthrax toxin receptor 2	-0.02	-0.03
OR	2HLE (A)	Ephrin type-B receptor 4	-0.01	0.04
OR	2HLE (B)	Ephrin-B2	-0.05	-0.04

Competing Interests

The authors declare that they have no competing financial, direct or indirect interests in connection with the work reported here.

References

- Arkin, M. and J. Wells, (2004). "Small-molecule inhibitors of protein-protein interactions: progressing towards the dream." *Nature Reviews Drug Discovery*, **3**(4): 301-317.
- Bogan, A. and K. Thorn, (1998). "Anatomy of hot spots in protein interfaces." *Journal of Molecular Biology*, **280**(1): 1-9.
- Chothia, C. (1976). "The nature of the accessible and buried surfaces in proteins." *Journal of Molecular Biology*, **105**(1): 1-12.
- Conte, L., Chothia, C. and J. Janin, (1999). "The atomic structure of protein-protein recognition sites." *Journal of Molecular Biology*, **285**(5): 2177-2198.
- Corbi-Verge, C. and P. Kim, (2016). "Motif mediated protein-protein interactions as drug targets." *Cell Communication and Signaling*, **14**(8): 1-12.
- Gutmanas, A., Alhroub, Y., Battle, G. M., Berrisford, J. M., Bochet, E., Conroy, M. J., Dana, J. M., Fernandez Montecelo, M. A., van Ginkel, G., Gore, S. P. and P. Haslam, (2013). "PDBe: Protein Data Bank in Europe." *Nucleic Acids Research*, **42**(1): 285-291.
- Hussain, A., Shanthi, V., Sheik, S., Jeyakanthan, J., Selvarani, P. and K. Sekar, (2002). "PDB Goodies – a web-based GUI to manipulate the Protein Data Bank file." *Acta Crystallographica Section D Biological Crystallography*, **58**(8): 1385-1386.
- Jones, S., and J. M. Thornton, (1996). "Principles of protein-protein interactions." *Proceedings of the National Academy of Sciences of the United States of America*, **93**(1): 13–20.
- Kastritis, P., Moal, I., Hwang, H., Weng, Z., Bates, P., Bonvin, A. and J. Janin, (2011). "A structure-based benchmark for protein-protein binding affinity." *Protein Science*, **20**(3): 482-491.
- Krissinel, E. and K. Henrick, (2007). "Inference of Macromolecular Assemblies from Crystalline State." *Journal of Molecular Biology*, **372**(3): 774-797.
- Laraia, L., McKenzie, G., Spring, D., Venkitaraman, A. and D. Huggins, (2015). "Overcoming Chemical, Biological, and Computational Challenges in the Development of Inhibitors Targeting Protein-Protein Interactions." *Chemistry & Biology*, **22**(6): 689-703.
- Lipinski, C., Lombardo, F., Dominy, B. and P. Feeney, (2001). "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings." *Advanced Drug Delivery Reviews*, **46**(1-3): 3-26.

Lewis, M. and D. Rees, (1985). "Fractal surfaces of proteins." *Science*, **230**: 1163 - 1165.

Miller, S., Janin, J., Lesk, A. and C. Chothia, (1987). "Interior and surface of monomeric proteins" *Journal of Molecular Biology*, **196**: 641 - 656.

Nero, T., Morton, C., Holien, J., Wielens, J. and M. Parker, (2014). "Oncogenic protein interfaces: small molecules, big challenges." *Nature Reviews Cancer*, **14**(4): 248-262.

Raj, M., Bullock, B. and P. Arora, (2013). "Plucking the high hanging fruit: A systematic approach for targeting protein-protein interactions." *Bioorganic & Medicinal Chemistry*, **21**(14): 4051-4057.

Thangudu, R., Bryant, S., Panchenko, A. and T. Madej, (2012). "Modulating Protein-Protein Interactions with Small Molecules: The Importance of Binding Hotspots." *Journal of Molecular Biology*, **415**(2): 443-453.

Toogood, P. (2002). "Inhibition of Protein-Protein Association by Small Molecules: Approaches and Progress." *Journal of Medicinal*

Chemistry, **45**(8): 1543-1558.

Tsodikov, O., Record, M. and Y. Sergeev, (2002). "Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature." *Journal of Computational Chemistry*, **23**(6): 600-609.

Vassilev, L. (2004). "In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2." *Science*, **303**(5659): 844-848.

Wells, J., and C. McClendon, (2007). "Reaching for high-hanging fruit in drug discovery at protein-protein interfaces." *Nature*, **450**(7172): 1001-1009.

White, A., Westwell, A. and G. Brahehi, (2008). "Protein-protein interactions as targets for small-molecule therapeutics in cancer." *Expert Reviews in Molecular Medicine*, **10**: e8. DOI: 10.1017/S1462399408000641.

This open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).