Exploring the charge space of protein–protein association: A proteomic study

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · AUGUST 2005

Impact Factor: 2.63 · DOI: 10.1002/prot.20489 · Source: PubMed

2 AUTHORS, INCLUDING:

Gideon Schreiber
Weizmann Institute of Science
134 PUBLICATIONS **7,186** CITATIONS

SEE PROFILE

Exploring the Charge Space of Protein–Protein Association: A Proteomic Study

Yossi Shaul and Gideon Schreiber*
The Weizmann Institute of Science, Rehovot, Israel

ABSTRACT **The rate of association of a protein complex is a function of an intrinsic basal rate and of the magnitude of electrostatic steering. In the present study we analyze the contribution of electrostatics towards the association rate of proteins in a database of 68 transient hetero-protein–protein complexes. Our calculations are based on an upgraded version of the computer algorithm *PARE*, which was shown to successfully predict the impact of mutations on k_{on} by calculating the difference in Columbic energy of interaction of a pair of proteins. *HyPare* (http://bip.weizmann.ac.il/*HyPare*), automatically calculates the impact of mutations on a per-residue basis for all residues of a protein–protein interaction, achieving a precision similar to that of *PARE*. Our calculations show that electrostatics play a marginal role (<10 fold) in determining the rate of association for about half of the complexes in the database. Strong electrostatic steering, which results in an increase of over 100-fold in k_{on}, was calculated for about 25% of the complexes. Applying *HyPare* to all 68 complexes in the database shows that a small number of residues are hotspots for association. About 40% of the hotspots are calculated to increase the rate of association upon mutation, and thus increase binding affinity. This is a much higher ratio than found for hotspots for dissociation, where the large majority cause weaker binding. About 40% of the hotspots are located outside the physical boundary of the binding site, making them ideal candidates for protein engineering. Our data shows that a majority of protein–protein complexes are not optimized for fast association. Hotspots are not evenly distributed between all types of amino acids. About 75% of all hotspots are of charged residues. This is understandable, as a charge-reverse mutant changes the total charge by 2. The small number of hydrophobic residues that are hotspots upon mutation probably relates to their location and surrounding. For 18 out of the 68 complexes in the database, experimental values of k_{on} are available. For these, a basal rate of association was calculated to be in the range of $10^4 M^{-1}s^{-1}$ to $10^7 M^{-1}s^{-1}$. Some of these rates were verified independently from experimental mutant data. The basal rates were correlated with the size of the proteins and the shape of the interface. Proteins 2005;60:341–352.** © 2005 Wiley-Liss, Inc.

Key words: electrostatic rate enhancement; protein complexes; protein–protein interactions; protein association; protein engineering; simulation; proteomics

INTRODUCTION

Proteins form complexes with other macromolecules on a fast time scale without sacrificing the specificity of interaction. This is true even in crowded solutions, where other macromolecules are added at large concentrations (mimicking the cellular environment).[1] The process of scanning a protein surface for a potential binding site is often (but not always) diffusion limited, and not a reaction limited process.[2] The association reaction begins by a random search of the two molecules for each other within the vast space of solution, followed by a precise docking of the two interfaces. Brownian motion dictates the rate of collision for a pair of globular proteins. According to the Smolochowski/Einstein equation, a rate of 10^9–$10^{10}\ M^{-1}s^{-1}$ is calculated independent on the exact size of the two molecules.[3] However, a random collision between two protein partners is far from assuring complex formation, as their relative orientations dictate additional constraints, which slow down the rate of the reaction by three to five orders of magnitude.[3,4] This rate can be substantially increased by favorable long-range electrostatic forces[3,5] or by favorable desolvation.[6] The role of electrostatic attraction in the process of association of protein complexes was studied successfully using Brownian dynamic simulations,[7,8] or by relating electrostatics to the average interaction potential between the active sites of two proteins.[9] Finally, a good quantitative relation between the electrostatic rate enhancement of protein complex formation and the electrostatic energy of interaction (ΔU) of a pair of proteins was demonstrated.[10–12] This relation was translated into a computer algorithm (*PARE*), which calculates the electrostatic component of the rate of association of a protein complex from the Columbic energy of interaction between two proteins (including the effect of screening of charges by salt).[11,13] The applicability of this program for mutant prediction and for protein design has

Grant sponsor: the Israel Science Foundation; Grant number: 389/02.

*Correspondence to: Gideon Schreiber, The Weizmann Institute of Science, 76100 Rehovot, Israel. E-mail: bcges@weizmann.ac.il

Received 14 October 2004; Revised 16 January 2005; Accepted 24 January 2005

Published online 10 May 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20489

been demonstrated.[11,12,14] Calculating the rate or association using a more detailed electrostatic model (solving the Poisson-Boltzmann equation as implemented in *DelPhi*[15] using the Parse 3 charge file[16]) was explored, but did not add to the quality of the results of the calculations.[17]

Experimental data suggest that mutations involving a charge change are those that most significantly alter the rate of association between proteins.[5,8,13,14,18,19] The magnitude of recorded change in k_{on} upon charge mutations varies greatly. In some cases a change of over 20-fold in k_{on} was recorded, while in others no change was observed. Hotspots for association are locations where the introduction of a mutation significantly affects k_{on}. In the current study we use a minimum threshold of greater than fourfold. The physical explanation for the location of hotspots relate to the charge distribution at their vicinity.[13] As electrostatic forces are long ranged, it is not surprising that some of the hotspots for association are located outside the physical binding site.

The proven success of *PARE* and the growing number of entries in the Protein Database Bank (PDB) provides an opportunity to examine the importance of electrostatics for association of protein complexes across a broader range of protein interactions. In this work we examined the electrostatic characteristics of 68 transient hetero-protein–protein complexes,[20] and how they are predicted to affect association. Our approach combines theoretical calculations with a solid backup of experimental evidence. The following questions will be addressed: how significant is the electrostatic contribution towards association in general, what is the magnitude of the basal rates of association (in the absence of electrostatics), how common are hotspots for association, and what their nature is. For the purpose of this study, we developed the computer algorithm *HyPare*, which automatically analyses protein–protein complexes on a per-residue basis for their contribution towards the rate of association.

RESULTS AND DISCUSSION

Electrostatic forces have a major impact on the stability of protein–protein interactions[21,22] and on their rate of association. We have previously developed the computer software *PARE* to calculate their contribution towards the rate of association from the Columbic energy of interaction between a pair of proteins.[11] The algorithm has proven to be highly accurate, both for validating the measured rates of existing mutagenesis data, and more importantly, for protein design of faster and tighter binding complexes.[11,12] In the current work we go one step further, by automating *PARE* to scan all amino acids of a protein complex for the expected rate change upon charge mutation, and to give a convenient graphical output of the data (*HyPare*: http://bip.weizmann.ac.il/HyPare/). We used *HyPare* to analyze a database of 68 different protein complexes for their respective electrostatic contribution to association and the prevalence of hotspots for association within these complexes.

Automating *HyPare*

HyPare was designed to introduce pseudocharges on all residues of a given protein complex to scan automatically all the residues of the proteins for expected changes in k_{on} upon mutation. The location of the pseudocharge on the amino acid is given in Table I. We tried to mimic the position of the pseudocharge to be as close as possible to that of a charge in a real charged residue. For some amino acids (like Gln, Asn, Leu, Val, etc.) the task is relatively easy (see Table I). For others (like Ala) the location of the pseudocharge is farther away from the location of a charge in a charged residue (Asp, Glu, Lys, or Arg). Still, we feel confident that at least to an approximation, this method should produce satisfactory results. It has been shown that the exact location of the charge is in many cases not critical.[13,14] Second, the exact location of the charge after mutation cannot be known *a priori*, as the precision of modeling the correct rotamer of surface side chains beyond Cβ is problematic.[23] The approach used in *HyPare* for charge allocation, and its precision in estimating the k_{on} values of mutant complexes, were evaluated for five different protein–protein interactions including 60 different mutations (Fig. 1). The experimental and *PARE* data shown in the figure were previously published.[11,12] Figure 1 clearly shows that *HyPare* performs almost as good as *PARE*, without the need to manually mutate the residues, and feed in new pdb files with modified coordinates into the calculations. Therefore, *HyPare* is a good high-throughput tool to investigate the role of electrostatics in protein complex association and as a first-estimate for protein design. Still, it may be useful to mutate the few hotspot positions identified by *HyPare in silico* followed by a step of minimization and recalculation using *PARE* to optimize the outcome. In most cases this last step will not cause any significant change in the results.

The Impact of Electrostatics on the Rate of Association

We have determined the *PARE* energy of interaction for 68 transiently binding hetero-protein complexes, taken from a database previously published.[20] The majority of complexes are electrostatically favored (Fig. 2 and Table II). Of those, most have a *PARE* energy of interaction of up to −10 kcal/mol at 0.15 M NaCl, with three complexes exhibiting extraordinary attraction. For a number of complexes, the electrostatics is calculated to be repulsive; however, the repulsion is always weak. For these complexes, the electrostatic forces slow down the rate of association by less than 10-fold. When removing two outliers (located 1.5 times the interquartile range, IQR) the distribution becomes normal (mean of −2.65 kcal/mol) with a p-value of 0.12 (Shapiro Wilk Normality test). Translating the electrostatic energy into a rate increase over a basal rate (which we define as the rate in the absence of electrostatics) shows that for about half of the analyzed protein complexes, electrostatic forces play only a moderate role in affecting association (for 30 complexes the change is less than 10-fold). For an additional 20 proteins, the rate increase is up to 100 fold, for eight proteins it is up to 1000 fold, and only for four proteins, the calculated rate increase is >10000 fold.

TABLE I. *HyPare* Charge Rules

Atom	Residue	Charge	Atom	Residue	Charge	Atom	Residue	Charge
OE1	GLU	−0.50	F3	ALF	−0.25	O2B	GTP	−0.50
OE2	GLU	−0.50	F4	ALF	−0.25	O1A	GTP	−0.50
OD1	ASP	−0.50	O1A	GDP	−0.50	O2A	GTP	−0.50
OD2	ASP	−0.50	O2A	GDP	−0.50	O1G	GTP	−0.67
ND1	HIS	0.00	O1B	GDP	−0.67	O2G	GTP	−0.67
NE2	HIS	0.00	O2B	GDP	−0.67	O3G	GTP	−0.67
NH1	ARG	0.50	O3B	GDP	−0.67	FE	HEM	2.00
NH2	ARG	0.50	O1G	GNP	−0.67	N_B	HEM	1.00
NZ	LYS	1.00	O2G	GNP	−0.67	N_D	HEM	1.00
CE2	TRP	0.00	O3G	GNP	−0.67	O1A	HEM	−0.50
CE3	TRP	0.00	O1B	GNP	−0.50	O2A	HEM	−0.50
OH	TYR	0.00	O2B	GNP	−0.50	O1D	HEM	−0.50
CZ	PHE	0.00	O1A	GNP	−0.50	O2D	HEM	−0.50
OE1	GLN	0.00	O2A	GNP	−0.50	NI	NI	2.00
NE2	GLN	0.00	CA	CA	2.00	FE	FEL	3.00
OD1	ASN	0.00	O1B	ADP	−1.00	O1	PO3	−1.00
ND2	ASN	0.00	O2B	ADP	−1.00	O2	PO3	−1.00
SG	CYS	0.00	O3B	ADP	−1.00	O2	PO3	−1.00
OG1	THR	0.00	O1A	ADP	−0.50	O1	CBM	−0.50
CG2	THR	0.00	O2A	ADP	−0.50	O2	CBM	−0.50
OG	SER	0.00	O1A	GSP	−0.50	CA	OC1	2.00
CD1	ILE	0.00	O2A	GSP	−0.50	O1	CBX	−0.50
CE	MET	0.00	O1B	GSP	−0.50	O2	CBX	−0.50
CD1	LEU	0.00	O2B	GSP	−0.50	CU	CU	2.00
CD2	LEU	0.00	S1G	GSP	−0.67	FE	HEC	2.00
CG1	VAL	0.00	O2G	GSP	−0.67	O1A	HEC	−0.50
CG2	VAL	0.00	O3G	GSP	−0.67	O2A	HEC	−0.50
CB	ALA	0.00	ZN	ZN	2.00	O1D	HEC	−0.50
OXT	*	−1.00	O4	PAL	−0.50	O2D	HEC	−0.50
AO1	FAD	−0.50	O5	PAL	−0.50	O1B	ATP	−0.50
A02	FAD	−0.50	O2	PAL	−0.50	O2B	ATP	−0.50
OP1	FAD	−0.50	O3	PAL	−0.50	O1A	ATP	−0.50
OP2	FAD	−0.50	O1P	PAL	−0.67	O2A	ATP	−0.50
FE	OFI	2.00	O2P	PAL	−0.67	O1G	ATP	−0.67
F1	ALF	−0.25	O3P	PAL	−0.67	O2G	ATP	−0.67
F2	ALF	−0.25	CL	CL	−1.00	O3G	ATP	−0.67
F2	ALF	−0.25	O1B	GTP	−0.50			

Charges are assigned to selected atoms according to the standard pdb atomic names. Here, the wild type charges are shown. Atoms of amino acids, including those with a charge of 0.00, are designated locations for the pseudo charge insertions (mutations). Atoms of heteroatom (e.g., ATP) are used for calculating ΔU but are not designated as locations of charge insertions. A charge of −1 is assigned to the OXT atom (oxygen terminus) if the coordinates are included in the pdb file. The N terminus is automatically assigned a +1 charge by the program. In some cases, the charge is distributed over two or more atoms, e.g., OE1 and OE2 of GLU and CE2 CE3 of TRP, where each atom potentially carries a charge of ±0.5 to a sum of ±1.

The Basal Rate of Association

Experimental values of k_{on} were determined for 18 out of the 68 complexes listed in our database (Table II). By subtracting the calculated electrostatic component of association from the experimental value of k_{on} it is possible to determine the basal rate of association. The calculated basal rates are in the range of less than $10^4\ M^{-1}s^{-1}$ to over $10^7\ M^{-1}s^{-1}$ [Fig. 3(A)]. The large variations in basal rates are seemingly independent of the electrostatic rate enhancement for the given complex. Thus, even for complexes where electrostatics play a minor role in dictating association, a large variation in basal rates is observed. The range of basal rates found in this study is somewhat larger than was estimated previously from the diffusion equations.[24] However, it fits the values obtained from mutant data, where the value of k_{on} of a theoretical mutant lacking electrostatic attraction is determined by extrapolating the available mutant data to $\Delta U = 0$ [Fig. 3(B)].[17] In four of the five complexes where mutant data is available (barnase-barstar, TEM-BLIP AChE-fas and Hiruding-Thrombin), a similar basal rate was obtained by both extrapolating the mutant k_{on} data to $\Delta U = 0$ and by subtracting ΔU from the rate of association [Fig. 3(A)]. Only for the association of Ras-Ral, no match was found between the two independent methods. The observed difference can be explained by the two-step mechanism of the association of Ras with Ral, with the first step being the formation of a measured intermediate along the association pathway.[12] Therefore, the association of Ras with Ral cannot be viewed as being purely diffusion limited, but

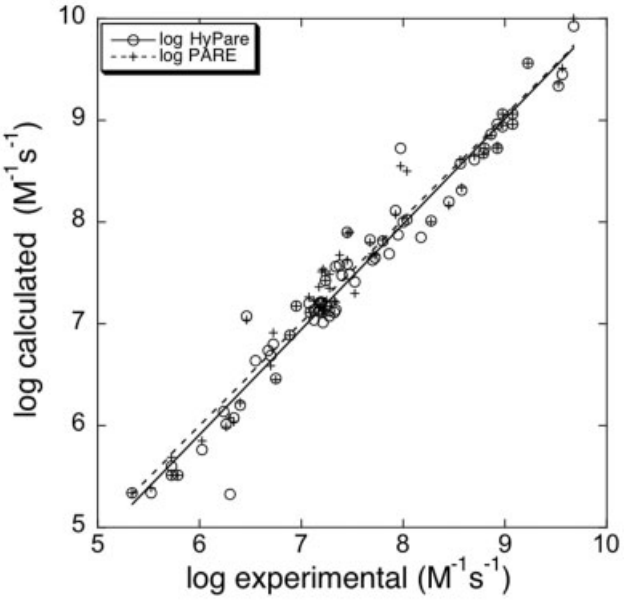


Fig. 1. Validating the calculated rates of association using *HyPare* against experimental data for five different protein complexes. Experimental wild-type and mutant data were plotted against calculated *HyPare* rates (+) using the charge rules given in Table I. For comparison, *PARE* prediction is also shown (○), using side-chain mutations followed by a step of minimization. The data shown in the figure are from the following complexes: barnase/barstar, TEM1/BLIP, Acetylcholinesterase/Fasciculin, Hirudin/Thrombin, and Ras/Ral. A correlation coefficient of 0.98 and a slope of 1 were determined between both calculations and the experimentally determined rates.

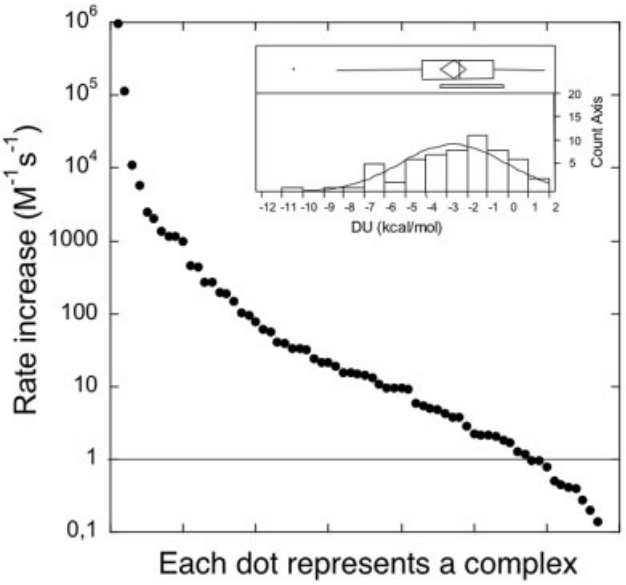


Fig. 2. Calculating the electrostatic rate-increase for 68 protein–protein complexes. The contribution of electrostatics to the rate of association was calculated using *HyPare* at 0.15 M NaCl [Eq. (5)] with the data and names of the protein complexes given in Table II. Each point on the graph represents a single protein complex. The points are sorted in increasing order. The insert shows the distribution of the Columbic energy of interaction as calculated by *HyPare* [Eq. (2)]. This distribution is normal (*p*-value of 0.12, Wilks Shapiro normality test) with a mean of −2.65 kcal/mol.

it is partially reaction limited. Indeed, the slope between the calculated and the experimental mutant data is not close to one (as it is in the other cases) but ~0.6 [Fig. 3(B)].

It is tempting to relate the basal rates of association with some physical character of the proteins shape and size. In Figure 3(C) we plot the calculated basal rates against $N/I \times \sqrt{S}$, where N is the number of atoms at the vicinity of the interface (atoms that are between 8 and 4 Å from the partner chain), I is the number of atoms at the interface, and S is the number of surface residues. The ratio N/I is a measure of the shape of the binding site (a round shape will give a smaller ratio than an oval shape), and $\sqrt{S}$ comes to normalize the data relative to the surface area of the proteins. The graph shows that the basal rate of association is faster for small proteins with round binding sites, with a coefficient of correlation of 0.61. The correlation to either the shape or the size alone is poorer (0.36, and 0.53, respectively). Interestingly, the basal rate determined for Ras/Ral by extrapolating ΔU to zero (blue dot) fits better our scheme than HyPare calculated basal rate. A different approach to explain the large variation in the basal rates for association could involve desolvation-mediated forces, or the nature of anchor residues in those protein-protein interactions.[6,25]

Abundance of Hotspots for Association

Mutant studies have shown that the contribution of most residues towards the rate of association is limited. However, some mutants do change k_{on} significantly. We call these locations hotspots for association. A good example for such is the Asp 163 Lys mutant located near the TEM1-BLIP interface, which causes an increase of 22 fold in k_{on}. Here we come to examine the prevalence of hotspots in protein complexes in general. Defining the threshold for which an expected change in k_{on} (upon mutation) can be considered as a "hotspot" is rather arbitrary. We choose to use two definitions: a change in k_{on} greater than fourfold (at $I = 0.15$) and >10-fold upon charge mutation. The first is most frequently used in the article, with the second coming to define locations of special importance towards association. *HyPare* identifies hotspots by assigning either a positive or a negative charge to each residue position using the charge definitions given in Table I. Figure 4 and Table II show the number of hotspots associated with each complex. There are 866 hotspots across the database, covering 63 of the 68 complexes. Although hotspots have a tendency to be more common in proteins with strong electrostatic attraction, the correlation between the two is weak (correlation coefficients of 0.7 for both fourfold and 10-fold change). Mutations at hotspot locations tend to cause a reduction in the rate of association more often than an increase (481 vs 381 for greater than fourfold). In 12 out of 68 complexes, no residue is anticipated to increase k_{on} over fourfold, and in half of the complexes no residue is anticipated to increase k_{on} >10-fold. Using this more stringent definition leaves 275 hotspots in the 68 complexes, from which only 69 increase k_{on} at this magnitude. These data suggest that although some proteins are

TABLE II. Summary of Calculated and Experimental Data for All Complexes Analyzed

pdb code		Chains 1st	Chains 2nd	ΔU	Δ rate	Experimental data k_{on} ($M^{-1}s^{-1}$)	Experimental data l_c (M)	Hotspots: Increase Total†	Hotspots: Increase 4 fold	Hotspots: Increase 10 fold	Hotspots: Decrease 4 fold	Hotspots: Decrease 10 fold
1a4y	RI / Ang[32]	A	B	−12.3	980090			79 (53)	51 (40)	15 (8)	28 (13)	13 (13)
1ai8	Hirudin / Thrombin[33]	H	I	−6.4	2690	1.03×10^8	0.15	20 (11)	4 (2)	1 (1)	11 (6)	4 (2)
1agr	Gl-α-1 / RGS4[34]	A	E	−1.2	4	1.7×10^6	0.14	11 (3)	4 (1)		7 (2)	2 (2)
1ak4	Cyc A / Hiv-1 capsid	A	D	−0.5	2							
1am4	P50-rhogap / Cdc42hs	A	D	−4.5	146			16 (7)	7 (4)		9 (3)	5 (5)
1aro	T7 ma pol / T7 lys	P	L	−2.3	13			11 (3)	3 (1)	1	8 (2)	4 (4)
1avw	Trypsin / Inhibitor[35]	A	B	−1.4	5	1.0×10^7	0.65	3			3	
1avz	NEF/ SH3 Domain	C	B	−1.2	4			6 (1)	3 (1)	1	3	2 (2)
1b6c	FKBP12 / Tgf-b	A	B	0.0	1							
1bi7	CDK6 / MTS1	A	B	−2.8	22			5	1		4	1 (1)
1bkd	H-ras / Son of sevenless-1	R	S	−3.1	32			8 (1)	2		6 (1)	3 (3)
1bml	Plasmin / Streptokinase[36]	A	C	−6.8	2062	5.4×10^7	0.16	18 (6)	6 (5)		12 (1)	5 (5)
1bmq	ICE / Inhibitor	A	B	−3.9	77			9 (3)	1 (1)		8 (2)	3 (3)
1bpl	α-1,4-glucan-4-glucan	A	B	0.8	0			9 (5)	4 (2)	1	5 (3)	
1brs	Barnase / Barstar[26]	A	D	−2.4	15	7.0×10^7		14 (5)	7 (4)	1 (1)	7 (1)	5 (5)
1bun	β2-bungarotoxin / Kunitz	A	B	0.6	1			6 (2)	4 (2)	1	2	2 (2)
1c0f	Gelsolin segment 1 / Actin[37]	S	A	0.8	0	2.1×10^4	0.16	8 (2)	6 (2)		2	
1cc0	Gtpase rhoa / Rho Gdi 1	A	E	0.0	1							
1cd9	G-csf / G-csf receptor	A	B	−4.1	93			7 (1)			7 (1)	1 (1)
1cgi	Serine protease / Inhibitor Ub yuh1-ubal / Ub yuh1-	E	I	−0.7	2			1 (1)			1 (1)	
1cmx	ubal	A	B	−3.7	61			19 (8)	10 (4)	1	9 (4)	6 (6)
1cqi	SCS-α / SCS-β	A	B	−2.4	14			9 (3)	1		8 (3)	1 (1)
1cse	Subtilisin carlsberg / Eglin-c	E	I	−0.5	2							
1cvs	FGF 2 /FGFR 1[38]	A	C	0.2	1	3.8×10^6	0.30	6 (3)	4(2)	1	2 (1)	1 (1)
1cxz	Rhoa / Pkn	A	B	−3.1	33			8 (0)	2	1	6	4 (4)
1d09	Aspartate carb / PALA	A	B	−5.0	265			19 (3)	8 (2)	2	11 (1)	5 (5)
1dhk	PPA / Inhibitor	A	B	1.8	0			16 (4)	13 (2)	2	3 (2)	0 (0)
1dn1	NSEC 1 / Syntaxin 1a	A	B	−4.1	101			16 (10)	3 (3)		13 (7)	2 (2)
1dtd	LCI / CPA2	A	B	−0.7	2			7 (1)	4 (1)	2	3	
1e96	RAC / P67PHOX	A	B	−1.5	5			7 (4)	3 (1)	1	4 (3)	1 (1)
1eay	CheY / CheA[39]	A	C	−3.3	39			11 (3)	5 (3)	1 (1)	6	5 (5)
1ef1	Moesin (ferm) / (tail)	A	C	−6.2	973			18 (6)	8 (2)	1	10 (4)	5 (5)
1eg9	NDO α/β subunits	A	B	−2.9	24			6 (1)	1		5 (1)	3 (3)
1emv	Im9 / Colicin E9[40]	A	B	−4.7	193	2.4×10^8	0.15	17 (7)	7 (2)		10 (5)	5 (5)
1eth	LIPF / Colipase	A	B	−2.0	9			4			4	1 (1)
1eui	ECUDG / UGI	A	C	−1.6	6			22 (12)	17 (8)	1 (1)	5 (4)	4 (4)
1f02	Intimin / Tir	I	T	−1.3	4			3 (1)	1		2 (1)	
1f60	eef1a / eef1ba	A	B	−6.3	1135			19 (9)	7 (3)	2 (1)	12 (6)	5 (5)
1fap	FKBP12 / Frap	A	B	−0.7	2			2 (1)			2 (1)	
1fcd	FCSD	A	C	−1.4	5			3 (1)			3 (1)	
1fin	CDK 2 / Cyclin a	A	B	−4.7	187			10 (1)	1		9 (1)	2 (2)
1frv	Hydrogenase	A	B	−7.0	2435			22 (4)	4	1	18 (4)	4 (4)
1fss	AChE / Fas II[41]	A	B	−5.5	453	2.7×10^7	0.07	20 (8)	8 (3)		12 (5)	5 (5)
1ggr	EIIA-GLC / HPR	A	B	−3.6	56			9 (4)	3 (2)		6 (2)	3 (3)
1hlu	Beta-actin / Profilin	A	P	−3.3	40			6 (0)			6	2 (2)

TABLE II. Continued

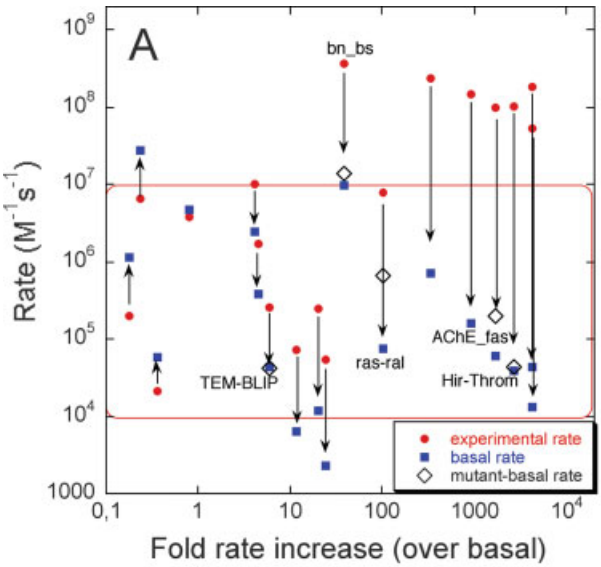
pdb code		Chains 1st	Chains 2nd	ΔU	Δ rate	Experimental data k_{on} ($M^{-1}s^{-1}$)	Experimental data l_c (M)	Hotspots: Increase Total†	Increase 4 fold	Increase 10 fold	Decrease 4 fold	Decrease 10 fold
1itb	IL-1 β / Type 1 IL-1R	A	B	−6.3	1148			13 (3)	3 (1)	1	10 (2)	6 (6)
1jtd	TEM / BLIP 2	A	B	0.7	0			11 (4)	8 (3)	3 (1)	3 (1)	1 (1)
1jtg	TEM / BLIP[11]	A	B	−1.2	11	2.6×10^5	0.17	6 (1)	2 (1)	1 (1)	4	2 (2)
1kac	knob / receptor[42]	A	B	−2.0	9	7.3×10^4	0.16	3			3	
1lfd	RalGDS / Ras[12]	A	B	−3.1	33	3.6×10^6	0.03	8 (1)	2		6 (1)	3 (3)
1noc	Oxidoreductase/ Transferase	A	B	−0.2	1			1			1	
1pdk	papD / papK	A	B	−2.5	16			4	1		3	
1qbk	Karyopherin β2 / Ran	B	C	−19.2	213205865			83 (55)	42 (30)	7 (3)	41 (25)	24 (24)
1rrp	Ran / Nup358[43]	A	B	−2.6	19	5.5×10^4	0.18	9 (1)	1	1	8 (1)	2 (2)
1smp	proteinase / Inhibitor	A	I	−0.7	2			6 (2)	1		5 (2)	1 (1)
1stf	Papain / StefinB[44]	E	I	−0.2	1							
1tba	TAFII230 / TBP	A	B	−8.3	10885			30 (15)	12 (7)		18 (8)	7 (7)
1tmq	α-amylase / Ragi inhibitor	A	B	−2.4	15			32 (15)	26 (12)	6 (1)	6 (3)	2 (2)
1uea	MMP-3 / TIMP-1[45]	A	B	1.4	0	2.0×10^5	0.18	1			1	
1wq1	Ras / Gap-334	R	G	−5.0	265			29 (12)	16 (8)	3	13 (4)	8 (8)
1ycs	P53 / 53bp2	A	B	−10.4	115527			49 (33)	28 (21)	5 (3)	21 (12)	13 (13)
1zbd	Rab-3a / Rabphilin-3a	A	B	−2.8	22			8 (4)	4 (2)	1	4 (2)	1 (1)
2kau	Lebsiella aerogenes urease	A	C	−0.9	3			2			2	
2kau	Lebsiella aerogenes urease	B	C	−2.0	9			4	1		3	1 (1)
2pcb	Ccp / Horse heart Cyt c	A	B	−7.7	5716			29 (22)	6 (5)	2 (1)	23 (17)	8 (8)
2pcf	Plastocyanin / Cyt f[46]	A	B	−6.5	1372	1.8×10^8	0.11	17 (3)	2 (1)	1	15 (2)	9 (9)
2sic	Subtilisin / Inhibitor[47]	E	I	1.2	0	6.5×10^6	0.14	2 (1)	2 (1)			
3eza	Enzyme i / Hpr[48]	A	B	−5.4	432	1.5×10^8	0.13	24 (10)	15 (8)	2	9 (2)	7 (7)
3hhr	HGH / Recpetor	A	B	−2.0	12	2.5×10^5	0.02	5			5	1 (1)

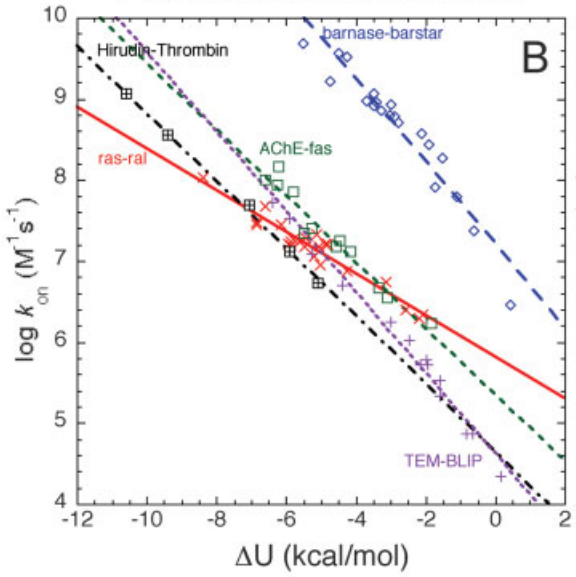
ΔU is the electrostatic energy of interaction computed at 0.150 M ionic strength (see methods for details). Δ rate is the rate increase or decrease imposed by ΔU at 0.150 M ionic strength.
†The number of hotspots located outside the interface is shown in parenthesis, calculated at 4 Å cutoff.

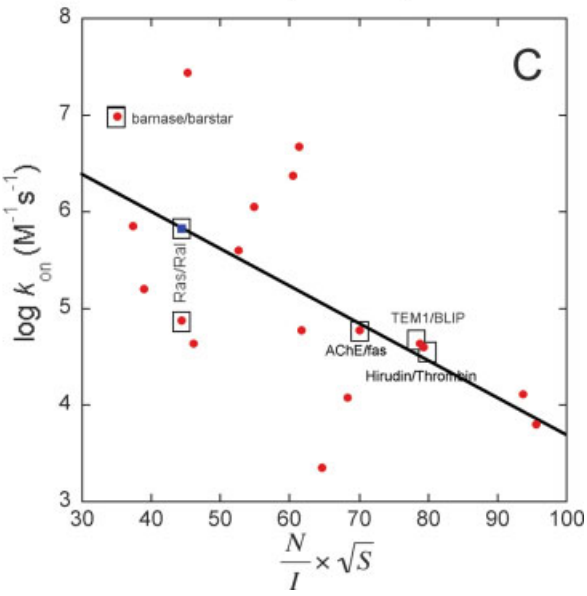
optimized for fast binding, many others are not. The reasons for this are quite clear. First, there is no evolutionary pressure for most protein complexes to be ultrafast binders. Second, electrostatic forces are long range and not specific. Thus, having too many proteins in the cell, which are strongly charged, may result in nonspecific binding, which could eventually slow down the overall yield of molecular recognition. Interestingly, physiological ionic strength provides sufficiently high-salt to mask nonspecific binding while still allowing for significant electrostatic steering of specific reactions.[26] For protein design, however, the lack of electrostatic optimization and our ability to calculate which residues should be mutated provide excellent opportunities to enhance association and binding. Of particular interest are residues located outside the physical binding site, as they can still increase association significantly without affecting the rate of dissociation.[11,12] These are found in 38 and 11 proteins for a k_{on} increase of greater than fourfold and 10-fold, respectively.

The Distribution of Hotspots in Protein Interfaces

Hotspots are found in protein complexes regardless of the magnitude of the electrostatic attraction between the two proteins (Fig. 4). A detailed analysis using *HyPare*, mutating one residue at a time along the protein chains of a complex, is shown in Figure 5 for two distinct pairs: Uracil–DNA–glycosylase binding its inhibitor [Fig. 5(A)], and cdk6 binding the multiple tumor suppressor [Fig. 5(B)]. The electrostatic contribution towards increasing the rate of association is 6 and 22 fold respectively (I = 150 mM) for these two complexes. While the electrostatic increase in k_{on} is moderate for both complexes, the electrostatic picture of their respective binding surfaces (middle panel), as well as their potential for protein engineering







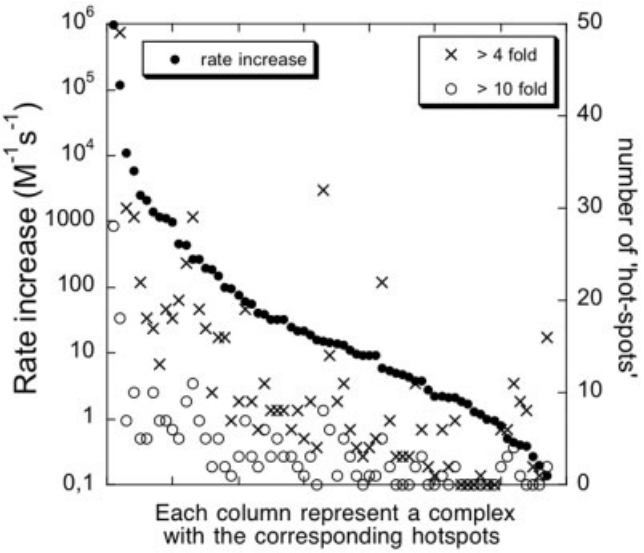


Fig. 4. The number of hotspots as a function of the electrostatic rate increase for 68 protein complexes. The change in rate is shown in each column with the corresponding number of hotspots determined at fourfold (×) and 10-fold (○) change upon charge mutation. A correlation coefficient of 0.70 and 0.72 was determined, respectively. The points are sorted in increasing order.

(left panel) is very different. Uracil–DNA–glycosylase shows a mixed electrostatic potential, while its inhibitor exhibits a strong negative potential. Consequently, all 22 hotspot positions in the complex are on Uracil–DNA–glycosylase, while none are located on the inhibitor. From the 22 hotspots, 17 are shown to increase k_{on} greater than fourfold, with the charge reverse mutant of Asp 64 increasing k_{on} >10-fold. The interaction of cdk6 binding the multiple tumor suppressor shows a fundamentally different electrostatic picture. Here, the electrostatic potential is mixed on both binding surfaces. Indeed, only five hotspots are identified, three on cdk6 and two on the suppressor. Four out of the five virtual mutations reduce binding and only one (the charge reverse of Asp 102 on cdk6) increases k_{on} greater than fourfold. Thus, this pair of proteins is not a good candidate for increasing k_{on} by protein engineering. The data shown in this figure (except the GRASP picture of the electrostatic potentials) are similar to those obtained using the convenient Web interface of *HyPare*.

Fig. 3. Basal rates of association. (**A**)The basal rate (red dots) was calculated by subtracting the electrostatic contribution (as calculated using *HyPare*) from the experimental value of k_{on} at the experimental ionic strength. For reference, the experimental rate is shown in blue. For five complexes, the basal rate was calculated also by extrapolating mutagenic data to $\Delta U = 0$ (diamonds), as shown in (**B**); for these complexes the basal rate is extrapolated from the fitted line at the energy value of 0 kcal/mol (no electrostatics contribution). (**C**) The basal rates from (**A**) are plotted against $N/I \times \sqrt{S}$, where N is the number of atoms at the vicinity of the interface (atoms that are between 8 and 4 Å from the partner chain), I is the number of atoms in the interface, and S is the number of surface residues. Two basal rates are given for Ras/Ral, the red dot using HyPare, and the blue dot is the mutant extrapolated one. A linear fit of the data gave a correlation coefficient of 0.61.

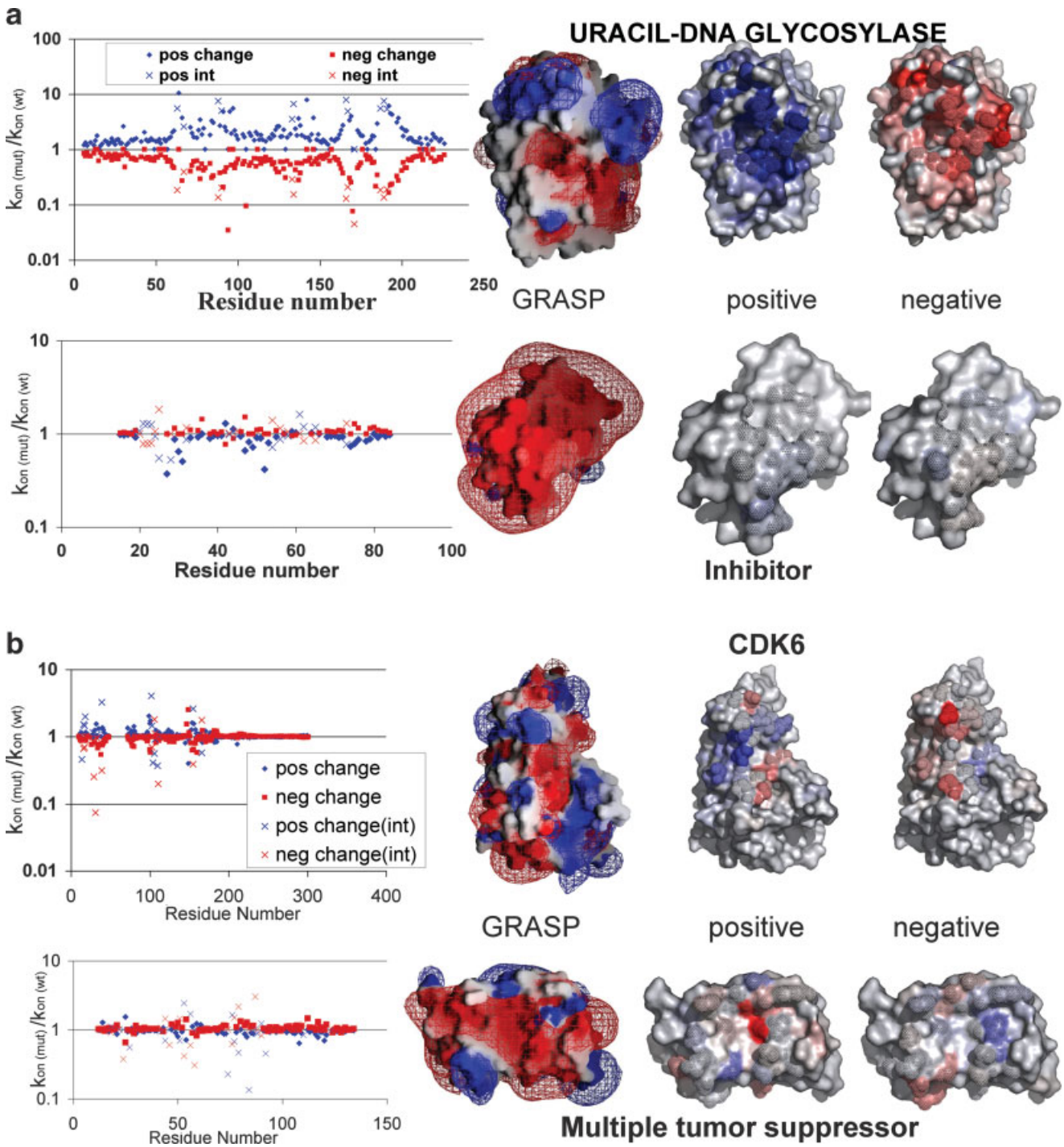


Fig. 5. Calculating the effect of a charge mutation along the protein sequence. The figure was produced using *HyPare* for (**a**). Uracil-DNA Glycosylase/Inhibitor (pdb code 1eui). (**b**) Cdk6/Multiple tumor suppressor (pdb code 1bi7). The graphs (left panel) show the potential of each residue position to alter the rate of association upon charge mutation. An × marks residues that are located within the binding site. The middle panel is a GRASP representation of the electrostatic potential of these proteins. The two right panels show the calculated rate changes upon mutation colored on the protein surfaces. The intensity of the blue (high) and red (low) colors indicates the rate change upon the introduction of a positive charge and negative charge, respectively. The interface is marked on the surface as dots. All calculations were done at an ionic strength of 0.15 M.

Charged Residues Are Dominant at Hotspot Locations

The amino acid propensities at hotspot locations are given in Figure 6. Not surprisingly, the charged residues, Arg, Asp, Glu, and Lys dominate the hotspot positions. A charge-reverse mutation of these residues changes the charge by 2, and not by 1 as it would have been in the case of adding a charge to a neutral residue. Changing the

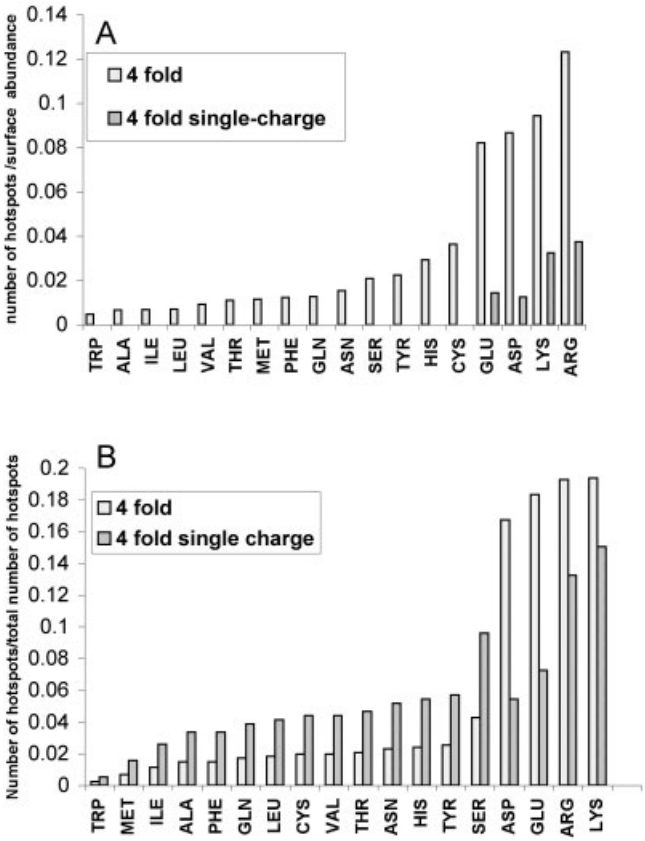


Fig. 6. Abundance of hotspots according to amino acid type. Hotspots are defined as residues that upon charge mutations alter the k_{on} by more than fourfold. For charged residues (which upon mutation change the charge by 2, e.g., Lys to Asp), Δk_{on} was also plotted to a single charge change (e.g., Lys to Ala). (**A**) The number of hotspots normalized relative to the abundance of the specific amino acid on the surface, while in (**B**) the number of hotspots is normalized relative to the total number of hotspots.

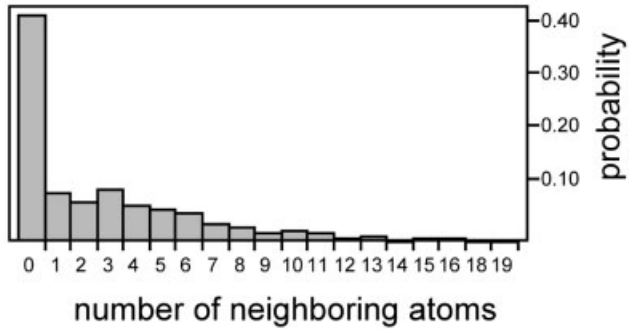


Fig. 7. The neighborhood of hotspots. A neighbor is considered as any atom of the other chain, which is within a distance of 4 Å. Thus, the x axis displays the number of neighbors, and the y axis display the fraction of hotspots that correspond to that number of neighbors. The total number of hotspots is 866.

charge by one for all residues decreases the count of hotspots of charged residues dramatically (Fig. 6). Yet, the charged residues are still dominant, with Serine turning out to be a primary hotspot source as well [Fig. 6(B)]. However, this can be attributed to the relative abundance of charged residues on the surface as seen from the number of hotspots normalized to surface abundance [Fig. 6(A)]. The same plot reveals another interesting observation, that polar surface residues (vs nonpolar surface residues) are more likely to be hotspots for association. This may be explained by the tendency of hydrophobic residues to be located in holes and not knobs on the surface, while polar residues tend to stick out.[20]

Hotspots Are Often Found at the Edges of the Interface

Columbic forces, which bring about the rate increase in association, act over a long range.[27] Therefore, it is not surprising that also residues located outside the binding site can strongly influence k_{on}.[11,12] The tendency of hotspots to be found outside the binding site is shown in Figure 7. About 40% of the hotspots have no neighboring atoms located on the partner chain at a distance closer than 4 Å, and about 15% of hotspots have no neighbors within a distance of 8 Å (data not shown). At a 4 Å definition of a binding site, these data demonstrate that hotspots outside the physical binding site are a common feature. Indeed, we have shown this to be not just theoretical, but true in practice, as demonstrated by introducing mutations outside the physical binding site.[11,12]

Three complexes have an extraordinary number of hotspots, Karyopherin α2/Ran, Ribonuclease inhibitor/Angiogenin, and P53/53bp2 (pdb codes 1qbk, 1a4y, and 1ycs, respectively). One possible explanation is to attribute the large number of hotspots to the relatively large surface area of these complexes (as found in 1qbk and 1a4y). However, the number of hotspots is not correlated with the size of the surface (correlation coefficient of 0.11; data not shown). Another characteristic of these three complexes is that they posses strongly polarized surfaces (positive vs negative). This, in turn, is the cause for the very high electrostatic energy of interaction calculated between the proteins within the complexes. It is difficult to know the precision of our calculations in these cases, and one may suspect that the simple mechanism of electrostatic-assisted association, which *PARE* and *HyPare* are based on, will not hold for these extreme cases.[13] Available kinetic mutant data for Ribonuclease Inhibitor binding angiogenin (pdb code 1a4y) shows that *HyPare* calculates the correct trend, but not the correct numbers (two mutations of charged residues, R5A and K40A, were shown to reduce k_{on} by about fourfold each,[28] but were calculated to reduce association by ~18-fold). A two-step mechanism for association, which results in a partially reaction controlled association, was shown to dominate the binding of this complex. We have already shown for the association between Ras and Ral that a two-step mechanism causes an overestimation of our calculated values.[12]

The work presented here deals only with the estimated changes in the rate of association of protein complexes. In the case that mutations are located outside the physical binding site, one can assume that these changes directly relate to the change in affinity, as long-range electrostatic forces have no effect on k_{off}.[11,12] Although our method seems to work equally good to predict changes in k_{on} for

charged mutations located within the binding site, these mutations may also cause large changes in the rate of dissociation. Changes in k_{off} are unpredictable using the here-described method, and have to be calculated using different algorithms.[22,29,30] However, the prediction of changes in k_{off} of short-range interactions are a much more difficult task compared to the one we took on ourself here, and the success of these methods is still lacking.

CONCLUSIONS

The current study came to give a broad picture of the role of electrostatics in determining association for many different protein complexes. The emerging picture is that electrostatics has a limited role in dictating association for most complexes, increasing their rate of association by up to 100-fold. Only in a few extreme cases, where association is functionally very important, electrostatic attraction has a major role. A somewhat surprising result was the large range of calculated basal rates of association, suggesting that the rate of association can be optimized not only through an electrostatic mechanism, but also by a second, yet unknown mechanism that results in faster binding. A good example for a protein interaction where such mechanism seems to be optimized is barnase-barstar (basal rate of 10^7 $M^{-1}s^{-1}$). The association of this complex is among the fastest recorded in the literature, yet electrostatic steering increases the rate by only 40-fold at $I = 22$ mM. Thus, the mechanism for the association of barnase-barstar may be structurally related to how easy it is to find the correct fit, through the binding of nonspecific hydrophobic surfaces or by preferential desolvation. More work has to be done to study this phenomenon. The function of most surface residues in promoting fast association is small. Yet, using *HyPare*, we identified many residues that do affect association by over fourfold. These are not equally distributed on all complexes, as some proteins do not have any hotspots at all, while others have many. Physically, hotspots for association are related to a certain local charge distribution of their near surrounding (found within 10 Å of the residue in question), previously shown using double-mutant cycles on the activated barnase-barstar complex.[31] The large charge distribution also explains why many hotspots are located outside the physical binding site. From a protein engineering perspective, the advantages are many, as it allows changing the binding affinity without disrupting the physical organization of the binding site. Many of the hotspots are charged residues, primarily due to their potential to add or reduce two charges at once. Taking this option away, they are not different from other polar residues. Moreover, polar residues have been found to be hotspots more often than nonpolar surface residues, possibly related to the higher degree of exposure of polar residues. Much of what is shown in this article can be done easily using the Web interface of *HyPare* (http://bip.weizmann.ac.il/*HyPare*), which was designed to give maximum information to the user.

MATERIALS AND METHODS

Database Construction

A database of 68 transient protein–protein heterodimers was constructed from the PDB.[20] These structures were first screened for nonsimilarity using BLAST, with a minimum p-value of 1×10^{-4}, followed by a second screen that examines the structural similarity of each possible pair, using the structural alignment Combinatorial Extension method (CE). For each pair that got a Z-score above 5, one of the proteins in the pair was removed from the DB. If the score was above 4, the complexes were aligned, and both were kept in the DB only if the interface location on the common monomer was different. The highest sequence identities according to the CE were 19.3%. The list of PDB codes is given in Table II.

HyPare Algorithm

1. Input
   1. Separate pdb file into chains.
   2. Combine desired chains (pairwise if complex has >2 chains).
   3. Assign charges to atoms (according to Table I).
   4. Find missing atoms (incomplete residues in the pdb file). Print to log file.
   5. Find interface residues (at least one atom = 4 Å from second chain).
   6. Reduce calculation set to charged atoms only.
   7. Calculate Coulombic energy of first chain [Eq. (1)].
   8. Calculate Coulombic energy of secondchain [Eq. (1)].
   9. Calculate Coulombic energy of the complex [Eq. (1)].
   10. Calculate energy difference [Eq. (2)].
   11. Calculate association rate at 0.022 M [Eq. (4)].
   12. Extract basal rate [Eq. (5)].
   13. calculate k_{on} at target salt [Eq. (6)].
   14. **If mode of work is finding hotspots [**
       For each residue position [
       For each charge type (Positive and Negative) [
       1. Assign charge to selected atoms.
       2. Calculate Coulombic energy of first chain [Eq. (1)].
       3. Calculate Coulombic energy of second chain [Eq. (1)].
       4. Calculate Coulombic energy of the complex [Eq. (1)].
       5. Calcuate energy difference [Eq. (2)].
       6. Calculate k_{on} at target salt [Eq. (6)].
       7. Print rate increase (the new rate/experimental rate).

       ]
       ]
       Else do next pdb (Do not find hotspots, just calculate ΔU)
       ]

The Coulombic energy, U, is calculated using:

$$U = \frac{1}{2}\sum_{i,j}\frac{q_i q_j}{4\pi\varepsilon_0\varepsilon r_{ij}}\frac{e^{-\kappa(r_{ij}-a)}}{1+\kappa a} \quad (1)$$

where i and j are the charged atoms in the proteins, ϵ is the dielectric constant set to 80, a is set to 6 Å, and κ is the Debye-Huckel screening parameter that relates to the ionic strength of the solution (for explanations, see ref. 11). The electrostatic energy of interaction is calculated from the difference between the electrostatic energy of the complex and the electrostatic energy of the two individual proteins:

$$\Delta U = U_{\text{complex}} - U_{\text{protein 1}} - U_{\text{protein 2}} \quad (2)$$

The relation between k_{on} and ΔU was calibrated using experimental data obtained from a number of different systems. The best fit was obtained when the calculations were done in a fixed ionic strength of 0.022 M, and a correction factor of 0.9 (which was determined from the best fit of the calculations to the experimental data for a number of different experimental systems) was used.[11] The association rate of a mutant protein at 0.022 M is calculated from Equation (3):

$$\ln k_{\text{on(mut)(I = 0.022)}} = (\Delta U_{\text{wt}} - \Delta U_{\text{mut}})/0.9 + \ln k_{\text{on(wt)}} \quad (3)$$

where $\ln k_{\text{on(wt)}}$ is determined experimentally at $I = 0.022$ M. In the case that the association rate is given at a different ionic strength is can be transformed using Equation (4):

$$\ln k_{\text{on(I = 0.022)}} = \ln k_{\text{on}} - \Delta U\left[\left(\frac{1}{1+\kappa a}\right)_{I = 0.022} - \left(\frac{1}{1+\kappa a}\right)_{1c}\right]/0.9 \quad (4)$$

where Ic is the ionic strength at which the experimental k_{on} was measured. The basal rate for association (in the absence of any electrostatic forces) is found using the relation:

$$\ln k^0_{\text{on}} = \Delta U_{\text{I = 0.022}}/0.9 + \ln k_{\text{on}(I = 0.022)} \quad (5)$$

The rate of association at any ionic strength Ic can be then calculated using:

$$\ln k_{\text{on(Ic)}} = \ln k^0_{\text{on}} - \Delta U\left[1 + \left(\frac{1}{1 + \kappa a}\right)_{Ic} - \left(\frac{1}{1 + \kappa a}\right)_{I = 0.022}\right]/0.9 \quad (6)$$

Although it is possible to calculate the basal rate and $k_{\text{on}(I\text{c})}$ directly from the experimental k_{on}, it is convenient to do the extra step calculation at 0.022 M as it fits the experimental procedure.

Charge Rules

To explore the charge space of protein complexes we introduced charge mutations into each residue position, which required assessing the exact location of the charges, as their coordinates are essential for calculating the Coulombic energy. One approach is to recalculate the coordinates using residue replacement algorithms. For instance, to introduce a positive charge into a location usually occupied by Alanine, one may replace it with Lysine, and recalculate the new coordinates in the PDB file. This process requires ample computer power if applied iteratively to each residue. In addition, algorithms that follow this approach introduce uncertainty while attempting to construct the correct rotamer.

HyPare applies a simplistic simulation approach by which the charge location was calculated at preferential atoms of the side chain of the residue in question. For instance, in the case of Alanine, we introduced a charge to the CB atom (see Table I), as if Alanine by nature carried that charge.

ACKNOWLEDGMENTS

We thank Jaime Prilusky from the Bioinformatics and Biological Computing for helping in creating the *HyPare* Web site. We thank Tal Peleg-Shulman for her critical reading of the manuscript.

REFERENCES

1. Minton AP. Implications of macromolecular crowding for protein assembly. Curr Opin Struct Biol 2000;10:34–39.
2. Schreiber G. Kinetic studies of protein–protein interactions. Curr Opin Struct Biol 2002;12:41–47.
3. Berg OG, von Hippel PH. Diffusion-controlled macromolecular interactions. Ann Rev Biophys Biophys Chem 1985;14:131–160.
4. Northrup SH, Erickson HP. Kinetics of protein–protein association explained by Brownian dynamic computer simulation. Proc Natl Acad Sci USA 1992;89:3338–3342.
5. Sheinerman FB, Norel R, Honig B. Electrostatic aspects of protein–protein interactions. Curr Opin Struct Biol 2000;10:153–159.
6. Camacho CJ, Kimura SR, DeLisi C, Vajda S. Kinetics of desolvation-mediated protein–protein binding. Biophys J 2000;78:1094–1105.
7. Elcock AH, Sept D, McCammon JA. Computer simulation of protein–protein interactions. J Phys Chem B 2001;105:1504–1518.
8. Gabdoulline RR, Wade RC. Protein–protein associaton: investigation of factors influencing asaociation rates by brownian dynamics simulations. J Mol Biol 2001;306:1139–1155.
9. Vijayakumar M, Wong KY, Schreiber G, Fersht AR, Szabo A, Zhou HX. Electrostatic enhancement of diffusion-controlled protein–protein association: comparison of theory and experiment on barnase and barstar. J Mol Biol 1998;278:1015–1024.
10. Kiel C, Serrano L, Herrmann C. A detailed thermodynamic analysis of ras/effector complex interfaces. J Mol Biol 2004;340:1039–1058.
11. Selzer T, Albeck S, Schreiber G. Rational design of faster associating and tighter binding protein complexes. Nat Struct Biol 2000;7:537–541.
12. Kiel C, Selzer T, Shaul Y, Schreiber G, Herrmann C. Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. Proc Natl Acad Sci USA 2004.
13. Selzer T, Schreiber G. New insight into the mechanism of protein–protein association. Proteins 2001;45:190–198.
14. Marvin JS, Lowman HB. Redesigning an antibody fragment for faster association with its antigen. Biochemistry 2003;42:7077–7083.
15. Gilson MK, Honig B. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. Proteins Struct Funct Genet 1988;4:7–18.
16. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 1994;98:1978–1988.
17. Selzer T, Schreiber G. Predicting the rate enhancement of protein complex formation from the electrostatic energy of interaction. J Mol Biol 1999;287:409–419.
18. Gabdoulline RR, Wade RC. Biomolecular diffusional association. Curr Opin Struct Biol 2002;12:204–213.
19. Zhou HX. Enhancement of protein–protein association rate by interaction potential: accuracy of prediction based on local Boltzmann factor. Biophys J 1997;73:2441–2445.

20. Neuvirth H, Raz R, Schreiber G. ProMate: a structure based prediction program to identify the location of protein–protein binding sites. J Mol Biol 2004;338:181–199.
21. Sheinerman FB, Honig B. On the role of electrostatic interactions in the design of protein–protein interfaces. J Mol Biol 2002;318: 161–177.
22. Lee LP, Tidor B. Barstar is electrostatically optimized for tight binding to barnase. Nat Struct Biol 2001;8:73–76.
23. Eyal E, Najmanovich R, Edelman M, Sobolev V. Protein side-chain rearrangement in regions of point mutations. Proteins 2003;50:272–282.
24. Schlosshauer M, Baker D. Realistic protein–protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. Protein Sci 2004;13:1660–1669.
25. Rajamani D, Thiel S, Vajda S, Camacho CJ. Anchor residues in protein–protein interactions. Proc Natl Acad Sci USA 2004;101: 11287–11292.
26. Schreiber G, Fersht AR. Rapid, electrostatic assisted, association of proteins. Nat Struct Biol 1996;3:427–431.
27. Nakamura H. Roles of electrostatic interaction in proteins. Q Rev Biophys 1996;29:1–90.
28. Shapiro R, Ruiz-Gutierrez M, Chen CZ. Analysis of the interactions of human ribonuclease inhibitor with angiogenin and ribonuclease A by mutagenesis: importance of inhibitor residues inside versus outside the C-terminal "hot spot." J Mol Biol 2000;302:497–519.
29. Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 2002;320:369–387.
30. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein–protein complexes. Proc Natl Acad Sci USA 2002;99:14116–14121.
31. Frisch C, Fersht AR, Schreiber G. Experimental assignment of the structure of the transition state for the association of barnase and barstar. J Mol Biol 2001;308:69–77.
32. Lee FS, Auld DS, Vallee BL. Tryptophan fluorescence as a probe of placental ribonuclease inhibitor binding to angiogenin. Biochemistry 1989;28:219–224.
33. Stone SR, Dennis S, Hofsteenge J. Quantitative evaluation of the contribution of ionic interactions to the formation of the Thrombin–Hirudin complex. Biochemistry 1989;28:6857–6863.
34. Lan KL, Zhong H, Nanamori M, Neubig RR. Rapid kinetics of regulator of G-protein signaling (RGS)-mediated Galphai and Galphao deactivation. Galpha specificity of RGS4 AND RGS7. J Biol Chem 2000;275:33497–33503.
35. Otlewski J, Zbyryt T, Dryjanski M, Bulaj G, Wilusz T. Single peptide bond hydrolysis/resynthesis in squash inhibitors of serine proteinases. 2. Limited proteolysis of *Curcurbita maxima* trypsin inhibitor I by pepsin. Biochemistry 1994;33:208–213.
36. Anonick PK, Wolf B, Gonias SL. Regulation of plasmin, miniplasmin, and streptokinase-plasmin complex by alpha 2-antiplasmin, alpha 2-macroglobulin, and antithrombin III in the presence of heparin. Thromb Res 1990;59:449–462.
37. Laham LE, Lamb JA, Allen PG, Janmey PA. Selective binding of gelsolin to actin monomers containing ADP. J Biol Chem 1993;268: 14202–14207.
38. Nugent MA, Edelman ER. Kinetics of basic fibroblast growth factor binding to its receptor and heparan sulfate proteoglycan: a mechanism for cooperactivity. Biochemistry 1992;31:8876–8883.
39. Silversmith RE, Appleby JL, Bourret RB. Catalytic mechanism of phosphorylation and dephosphorylation of CheY: kinetic characterization of imidazole phosphates as phosphodonors and the role of acid catalysis. Biochemistry 1997;36:14965–14974.
40. Wallis R, Moore GR, James R, Kleanthous C. Protein–protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. Biochemistry 1995;34:13743–13750.
41. Eastman J, Wilson EJ, Cervenansky C, Rosenberry TL. Fasciculin 2 binds to the peripheral site on acetylcholinesterase and inhibits substrate hydrolysis by slowing a step involving proton transfer during enzyme acylation. J Biol Chem 1995;270:19694–19701.
42. Kirby I, et al. Adenovirus type 9 fiber knob binds to the coxsackie B virus-adenovirus receptor (CAR) with lower affinity than fiber knobs of other CAR-binding adenovirus serotypes. J Virol 2001;75: 7210–7214.
43. Villa Braslavsky CI, Nowak C, Gorlich D, Wittinghofer A, Kuhlmann J. Different structural and kinetic requirements for the interaction of Ran with the Ran-binding domains from RanBP2 and importin-beta. Biochemistry 2000;39:11629–11639.
44. Abrahamson M, Barrett AJ, Salvesen G, Grubb A. Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. J Biol Chem 1986;261:11282–11289.
45. Troeberg L, et al. *E. coli* expression of TIMP-4 and comparative kinetic studies with TIMP-1 and TIMP-2: insights into the interactions of TIMPs and matrix metalloproteinase 2 (gelatinase A). Biochemistry 2002;41:15025–15035.
46. Schlarb-Ridley BG, Bendall DS, Howe CJ. Role of electrostatics in the interaction between cytochrome f and plastocyanin of the cyanobacterium *Phormidium laminosum*. Biochemistry 2002;41: 3279–3285.
47. Masuda-Momma K, et al. Interaction of subtilisin BPN′ and recombinant Streptomyces subtilisin inhibitors with substituted P1 site residues. J Biochem (Tokyo) 1993;114:553–559.
48. Garrett DS, Seok YJ, Peterkofsky A, Clore GM, Gronenborn AM. Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the *Escherichia coli* phosphotransferase system. Biochemistry 1997;36:4393–4398.