Rationalizing 5000-Fold Differences in Receptor-Binding Rate Constants of Four Cytokines

Xiaodong Pang, Sanbo Qin, and Huan-Xiang Zhou*
Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida

ABSTRACT The four cytokines erythropoietin (EPO), interleukin-4 (IL4), human growth hormone (hGH), and prolactin (PRL) all form four-helix bundles and bind to type I cytokine receptors. However, their receptor-binding rate constants span a 5000-fold range. Here, we quantitatively rationalize these vast differences in rate constants by our transient-complex theory for protein-protein association. In the transient complex, the two proteins have near-native separation and relative orientation, but have yet to form the short-range specific interactions of the native complex. The theory predicts the association rate constant as $k_a = k_{a0} \exp(-\Delta G^*_{el}/k_B T)$ where k_{a0} is the basal rate constant for reaching the transient complex by random diffusion, and the Boltzmann factor captures the rate enhancement due to electrostatic attraction. We found that the vast differences in receptor-binding rate constants of the four cytokines arise mostly from the differences in charge complementarity among the four cytokine-receptor complexes. The basal rate constants (k_{a0}) of EPO, IL4, hGH, and PRL were similar (5.2×10^5 $M^{-1}s^{-1}$, 2.4×10^5 $M^{-1}s^{-1}$, 1.7×10^5 $M^{-1}s^{-1}$, and 1.7×10^5 $M^{-1}s^{-1}$, respectively). However, the average electrostatic free energies (ΔG^*_{el}) were very different (−4.2 kcal/mol, −2.4 kcal/mol, −0.1 kcal/mol, and −0.5 kcal/mol, respectively, at ionic strength = 160 mM). The receptor-binding rate constants predicted without adjusting any parameters, 6.2×10^8 $M^{-1}s^{-1}$, 1.3×10^7 $M^{-1}s^{-1}$, 2.0×10^5 $M^{-1}s^{-1}$, and 7.6×10^4 $M^{-1}s^{-1}$, respectively, for EPO, IL4, hGH, and PRL agree well with experimental results. We uncover that these diverse rate constants are anticorrelated with the circulation concentrations of the cytokines, with the resulting cytokine-receptor binding rates very close to the limits set by the half-lives of the receptors, suggesting that these binding rates are functionally relevant and perhaps evolutionarily tuned. Our calculations also reproduced well-observed effects of mutations and ionic strength on the rate constants and produced a set of mutations on the complex of hGH with its receptor that putatively enhances the rate constant by nearly 100-fold through increasing charge complementarity. To quantify charge complementarity, we propose a simple index based on the charge distribution within the binding interface, which shows good correlation with ΔG^*_{el}. Together these results suggest that protein charges can be manipulated to tune k_a and control biological function.

INTRODUCTION

Cytokines are a large family of small proteins that, by binding to specific cell surface receptors, initiate signals critical for cell proliferation, differentiation, and apoptosis. Class-I helical cytokines form four-helix bundles. Erythropoietin (EPO), interleukin-4 (IL4), human growth hormone (hGH), and prolactin (PRL) are well-known members of this class. EPO is produced mainly in the adult kidney, and is responsible for red blood cell production and maintenance. IL4 is produced primarily by activated T-cells and mast cells, and is involved in stimulation of activated B-cells, proliferation of T-cells, and differentiation of $CD4^+$ T-cells into Th2 cells. The hGH is produced mainly in the anterior pituitary gland; its function is to stimulate growth, cell reproduction, and regeneration. PRL is also produced in the anterior pituitary gland; it simulates the mammary glands to produce milk and has other functions including an essential role in the maintenance of immune system functions.

To elucidate signaling mechanisms and develop therapeutic applications, the interactions of these four cytokines with their receptors have been intensively studied in the past two decades (1–18). All of them bind to type I cytokine receptors. Each of these cytokines has two receptor binding sites, referred to as site 1 and site 2. Initial binding to a receptor molecule via the high-affinity site (Fig. 1, *A–D*) primes the subsequent binding of a second receptor molecule to the remaining site. The resulting dimerization of the receptor molecules initiates the signaling cascade. Despite these similarities, the association rate constants of EPO, IL4, hGH, and PRL with their first receptor molecules, EPO receptor (EPOR), IL4 alpha receptor (IL4Rα), hGH receptor (hGHR), and PRL receptor (PRLR), vary widely: 4.0×10^8 $M^{-1}s^{-1}$ (6), 1.3×10^7 $M^{-1}s^{-1}$ (8), 3.2×10^5 $M^{-1}s^{-1}$ (15), and 8.0×10^4 $M^{-1}s^{-1}$ (17), respectively (at an ionic strength of 160 mM). Here, we use our recently developed transient-complex theory for protein-protein association (19–22) to quantitatively rationalize the 5000-fold differences in receptor-binding rate constants among the four cytokines, and highlight the role of protein charges in controlling biological function via tuning association rates.

The transient complex refers to the intermediate along the binding pathway, in which the two associating proteins have near-native separation and relative orientation, but have yet

Submitted April 1, 2011, and accepted for publication June 8, 2011.

*Correspondence: hzhou4@fsu.edu

Editor: Gerhard Hummer.

0006-3495/11/09/1175/9 $2.00

doi: 10.1016/j.bpj.2011.06.056

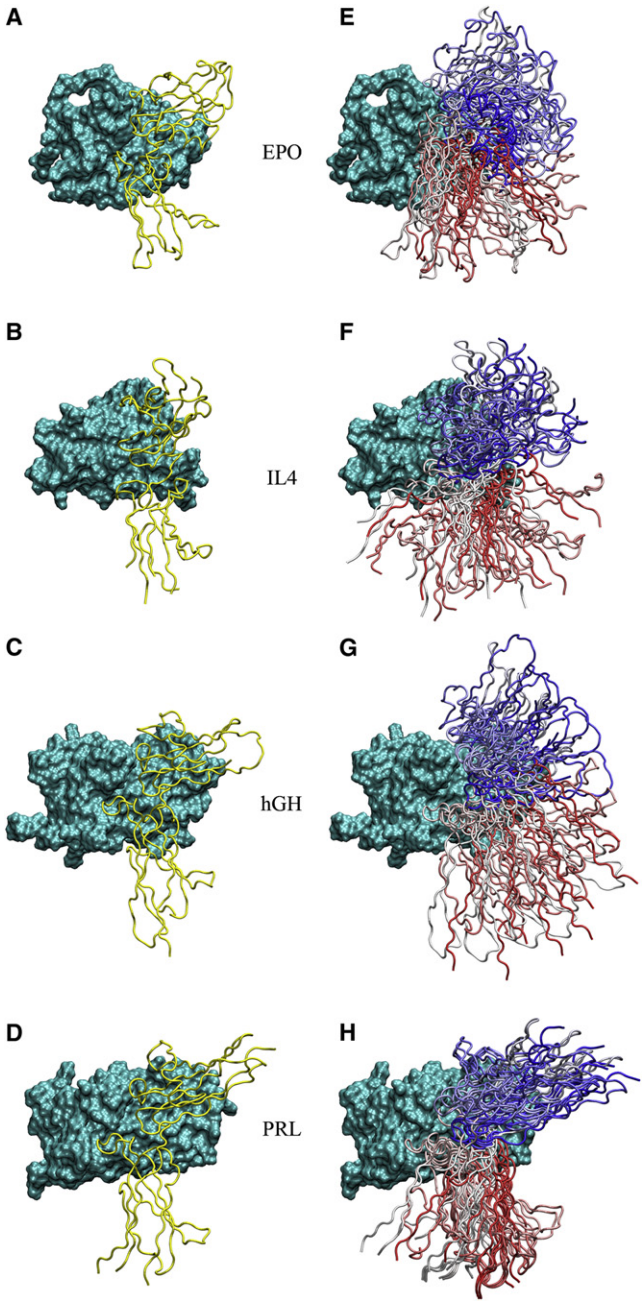

FIGURE 1 Structures of the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR high-affinity complexes. (*A*) The EPO-EPOR native complex. EPO is shown as surface and EPOR as tube. (*B*) The IL4-IL4Rα native complex. (*C*) The hGH-hGHR native complex. (*D*) The PRL-PRLR native complex. (*E*) The EPO-EPOR transient-complex ensemble, as illustrated by eight representative poses. In each pose EPOR is represented as a tube with color varying from blue at the N-terminal to red at the C-terminal. (*F*) The IL4-IL4Rα transient-complex ensemble. (*G*) The hGH-hGHR transient-complex ensemble. (*H*) The PRL-PRLR transient-complex ensemble.

to form the short-range specific interactions of the native complex (20). The transient-complex theory predicts the association rate constant as

$$k_a = k_{a0} \exp\left(-\Delta G^*_{el}/k_B T\right), \quad (1)$$

where k_{a0} is the basal rate constant for reaching the transient complex by random diffusion, $-\Delta G^*_{el}$ is the average electrostatic interaction free energy of the transient complex, k_B is the Boltzmann constant, and T is the absolute temperature. Electrostatic attraction enhances the association rate by increasing the probability of reaching the transient complex. Equation 1 allows us to isolate the electrostatic contribution to k_a, via ΔG^*_{el}, from the basal rate constant k_{a0}. Our implementation of Eq. 1 is fully automated and free of adjustable parameters.

EPO natively contains three N-linked glycans at positions N24, N38, and N83 and one O-linked glycan at S126. Previous studies demonstrated that desialylation of EPO results in an increased EPOR affinity (2). The increase in affinity is primarily due to an increase in the association rate constant k_a, with very little change in the dissociation rate constant k_d. Introduction of a single negative charge on the protein surface has an effect similar to that of glycosylation (6), suggesting that electrostatic interactions play a critical role in EPO-EPOR binding. Our calculations based on Eq. 1 found a significant electrostatic contribution to the EPO-EPOR binding rate, enhancing the k_a of nonglycosylated EPO (NGE) by 1175-fold at ionic strength = 160 mM. The rate enhancement arose from strong electrostatic complementarity between EPO and EPOR, which feature highly positive and negative electrostatic surfaces, respectively. The calculated k_a results agree well with experimental data for NGE and two variants with the four glycosylation sites replaced either by negatively charged Glu or by positively charge Lys, over a wide range of ionic strength.

Compared to EPO-EPOR binding, our calculations found that electrostatic contributions progressively weaken for IL4-IL4Rα and hGH-hGHR binding, and even become mildly unfavorable for PRL-PRLR binding. In line with these calculations, the electrostatic surfaces of IL4 and IL4Rα show moderate complementarity, whereas those of hGH and hGHR and of PRL and PRLR are largely mixed. Therefore, the vast differences in association rate constants of the four cytokine-receptor complexes are simply dictated by the degree of electrostatic complementarity. To simplify the calculation of this complementarity, we propose an index based on the charge distribution within the binding interface. Although a previous attempt by McCoy et al. (23) using a charge distribution-based index to represent electrostatic complementarity was unsuccessful, our charge complementarity index shows good correlation with ΔG^*_{el}.

Charge mutations at numerous positions on both hGH and hGHR were found to have minimal effect on k_a (13–15). Given that electrostatic interactions contribute little to the hGH-hGHR association rate; this finding is now easy to understand. Our calculations here on these mutations found that they indeed do not significantly affect k_a. However, the results on the EPO-EPOR binding rate suggest that significant electrostatic rate enhancement can be introduced to hGH-hGHR binding through mutation. We found that

concurrent charge reversal at two positions in the hGHR binding site on hGH, D171R and E174R, and at two positions in the hGH binding site on hGHR, R43E and R217E, significantly increase charge complementarity, and putatively increase the hGH-hGHR binding rate constant by nearly 100-fold.

Cellular processes are always faced with competing pathways and are thus likely under kinetic control rather than thermodynamic control (22). In particular, a given cytokine receptor may bind its cytokine but may also be lost due to degradation or other reasons. The receptor is useful only if cytokine binding occurs first. The half-life of the receptor therefore sets a lower limit for the pseudo-first order receptor-binding rate constant, which is the product of the bimolecular binding rate constant k_a and the cytokine circulation concentration. We find that the vastly different k_a values of the four cytokines studied here are compensated by vastly different circulation concentrations, such that the pseudo-first order rate constants are all close to the limits set by the half-lives of the receptors. Therefore, the cytokine-receptor binding rates appear to be evolutionarily tuned to ensure that all the receptors produced would participate in cytokine binding.

THEORETICAL METHODS

Starting from the structure of the native complex, our implementation of the transient-complex theory for calculating the protein-protein association rate constant is fully automated and free of adjustable parameters (20,21). It consists of three main components: i), generation of the transient-complex ensemble; ii), calculation of the basal rate constant; and iii), calculation of the electrostatic interaction free energy $-\Delta G^*_{el}$. We now briefly describe the preparation of the native complex and the components of the transient-complex theory.

Structure preparation for native complexes

The structures of the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR native complexes were from Protein Data Bank entries 1EER (4), 1IAR (9), 1A22 (14), and 3NPZ (18), respectively. Three of these complexes consist of the receptor extracellular domain bound to site 1 of the cytokine, but the IL4-IL4Rα complex involves site 2 instead (Fig. 1, *A–D*). Two missing loops in hGHR were built by Modeler (24). All hydrogen atoms were added and energy minimized by the AMBER program.

In addition to the wild-type complexes, a large number of mutations were studied. These include EPO NGE mutants in which the four glycosylated sites were replaced either by Glu or by Lys. These mutants were referred to as NGE-Glu and NGE-Lys, respectively. Four charge reversal mutations of IL4 (K77E, R81E, K84D, and R85E) and a large number of charge mutations on the hGH-hGHR complex for which experimental k_a data are available (8,13–15) were also studied. Finally, a variant hGH-hGHR complex, containing double mutations D171R and E174R on hGH and double mutations R43E and R217E on hGHR, was designed to increase k_a. The replaced side chains were optimized by energy minimization.

Generation of transient-complex ensembles

The transient complex was identified with the outer boundary of the bound-state energy well (19), after generating the interaction energy landscape around the native complex in the six-dimensional space of relative translation and relative rotation. For each complex, the cytokine subunit was fixed in space. The relative translation of the receptor subunit was represented by the displacement vector $\mathbf{r}$, and the relative rotation was represented by a body-fixed unit vector $\mathbf{e}$ and a rotation angle χ around the vector. The native complex has $r \equiv |\mathbf{r}| = 0$ and $\chi = 0$.

The short-range interaction energy around the native complex was represented by the total number N_c of contacts between two lists of representative atoms across the binding interface. In general, the value of N_c decreases as the two subunits move apart; along the way the range of allowed χ values (i.e., those corresponding to clash-free poses) exhibits a sharp increase. The value of N_c at the midpoint of this sharp transition, denoted as N^*_c, defined the transient complex. That is, the transient-complex ensemble consisted of all the poses with $N = N^*_c$. The values of N_c were 44, 32, 61, and 58 for the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR native complexes, respectively. Using a total of 8×10^6 poses each for the four systems, the corresponding values of N^*_c defining the transient complexes were determined to be 18, 13, 19, and 19, respectively.

Calculation of basal rate constants by force-free Brownian dynamics simulations

The basal rate of constant for reaching the transient complex by random diffusion was obtained from Brownian dynamics simulations as previously described (20). The translational diffusion constants of the proteins were assigned according to their molecular mass (25). These were 10.3 Å^2/ns and 9.4 Å^2/ns for EPO and EPOR; 10.9 Å^2/ns and 9.8 Å^2/ns for IL4 and IL4Rα; 9.7 Å^2/ns and 9.4 Å^2/ns for hGH and hGHR; and 9.8 Å^2/ns and 9.4 Å^2/ns for PRL and PRLR. For each cytokine-receptor complex, 4000 Brownian dynamics trajectories were used to calculate k_{a0}.

Calculation of ΔG^*_{el}

As in previous studies (20,21,26), 100 poses from the transient-complex ensemble were randomly selected to calculate ΔG^*_{el}. For each pose, the electrostatic interaction free energy, ΔG_{el} was calculated as (20,21,27)

$$\Delta G_{el} = G_{el}(\text{complex}) - G_{el}(\text{cytokine}) - G_{el}(\text{receptor}), \quad (2)$$

where G_{el} is the total electrostatic free energy of a solute molecule. Here, complex refers to a pose from the transient-complex ensemble. The average of ΔG_{el} over the 100 poses produced ΔG^*_{el}.

Electrostatic calculations were done by the Adoptive Poisson-Boltzmann Solver (APBS version 1.2) (28), with AMBER charges (29) and Bondi radii (30). The full, nonlinear Poisson-Boltzmann equation was solved. The dielectric constant of the solute molecule was set to 4, and the dielectric constant of the solvent was set to 78.5, corresponding to a temperature of 300 K. Atomic charges were mapped to grid points with the cubic B-spline discretization, with the *chgm* flag set to *spl2*. Following our previous studies on protein-protein and protein-RNA association (20,21,27), the dielectric boundary was specified as the van der Waals surface by setting the *srfm* flag to *mol* and *srad* to 0. The range of ionic strength (I) was from 60 to 1010 mM.

Each APBS calculation started with a coarse grid with dimensions of $193 \times 193 \times 193$ covering a volume of 288 Å $\times$ 288 Å $\times$ 288 Å around the solute molecule with the single Debye-Hückel boundary condition. The Poisson-Boltzmann equation was then solved on a fine grid with dimensions of $193 \times 193 \times 193$ covering a volume of 144 Å $\times$ 144 Å $\times$ 144 Å centered on the binding interface of the complex.

Quantification of charge complementarity

Our charge complemtarity index (CCI) was based on the charge distribution within the binding interface. Charged residues were represented by one or

two side-chain atoms: OD1 and OD2 for Asp, OE1 and OE2 for Glu; NZ for Lys, and NH1 and NH2 for Arg. Charged atom pairs across the binding interface of the native complex were collected with a 7-Å cutoff. These pairs were grouped into four kinds: +−; −+; ++; and − −. The charge complementarity index was then calculated as

$$\mathrm{CCI} = \left| \sum_i \frac{1}{d_i^{+-}} - \sum_i \frac{1}{d_i^{-+}} \right| - \left(\sum_i \frac{1}{d_i^{++}} + \sum_i \frac{1}{d_i^{--}} \right), \tag{3}$$

where d_i^{+-}, d_i^{-+}, d_i^{++}, and d_i^{--} denote atom-atom distances of the charged pairs. The first term favors opposite-charge pairs going in one direction, whereas the second term penalizes like-charge pairs. A protein complex with many positive charges on one side of the interface and many negative charges on the other has a high CCI score; mixed charges on either side result in a low score; and the presence of charges with the same sign on both sides of the interface results in a negative score.

RESULTS

The focus of this study is the rate constants for forming the high-affinity complexes of EPO, IL4, hGH, and PRL with their respective receptors. The four cytokines all form four-helix bundles, with C_α pairwise root mean-square deviations (RMSD) ~3.5 Å. The four receptors each consist of two fibronectin-III domains and are also structurally similar, with RMSD again ~3.5 Å. The EPO-EPOR, hGH-hGHR, and PRL-PRLR complexes superimpose to RMSD ~4 Å, but the IL4-IL4Rα complex involves a different binding site such that the orientation of IL4 is flipped (Fig. 1, *A–D*). Despite the structural similarities, the association rate constants of the four systems differ by 5000-fold (6,8,15,17). We now quantitatively rationalize these vast differences in k_a.

Transient complexes of the EPO-EPOR, IL4-IL4Rα, hGH-HGHR, and PRL-PRLR pairs

Out of 8×10^6 poses each sampled around the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR native complexes, the transient complexes were determined. The four transient-complex ensembles contained 21,317, 10,351, 28,720, and 7410 poses, respectively. Like the native complexes, the transient complexes of these cytokine-receptor systems are also similar (again the orientation of IL4 is flipped). This is not surprising, because the transient complexes are determined by the native complexes. In Fig. 1, *E–H*, we display eight representative poses each from the four transient complexes. The averages and standard deviations of cytokine-receptor separations in the four transient complexes were 6.1 ± 1.2 Å, 5.3 ± 0.9 Å, 6.1 ± 1.2 Å, and 5.5 ± 1.1 Å. The relative rotations of the subunits in the four transient complexes were also similar. Relative to the cytokines, the body-fixed unit vectors of the receptors were mostly restricted to cones spanning 30°, 23°, 20°, and 30°, respectively. The averages and standard deviations of the rotation angles around the body-fixed vectors were −7° ± 15°, 6° ± 17°, 17° ± 10°, and −6° ± 13°, respectively.

Predicted binding rate constants at $I = 160$ mM

The receptor-binding rate constants of EPO, IL4, hGH, and PRL obtained by the transient-complex theory were 6.2×10^8 $M^{-1}s^{-1}$, 1.3×10^7 $M^{-1}s^{-1}$, 2.0×10^5 $M^{-1}s^{-1}$, and 7.6×10^4 $M^{-1}s^{-1}$, respectively, at ionic strength = 160 mM. These compare favorably with the experimental values, 4.0×10^8 $M^{-1}s^{-1}$ (6), 1.3×10^7 $M^{-1}s^{-1}$ (8), 3.2×10^5 $M^{-1}s^{-1}$ (15), and 8.0×10^4 $M^{-1}s^{-1}$ (17), respectively. Our calculations thus rationalize the 5000-fold differences in k_a in the four cytokine-receptor systems.

What accounts for the vast differences in k_a? We found the basal rate constants to be similar, 5.2×10^5 $M^{-1}s^{-1}$, 2.4×10^5 $M^{-1}s^{-1}$, 1.7×10^5 $M^{-1}s^{-1}$, and 1.7×10^5 $M^{-1}s^{-1}$, respectively, for the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR pairs. However, the values of the average electrostatic free energy ΔG_{el}^* were very different, −4.2 kcal/mol, −2.4 kcal/mol, −0.1 kcal/mol, and 0.5 kcal/mol, respectively, for the four systems. Therefore, electrostatic interactions significantly enhance the receptor binding rates of EPO and IL4 (1175-fold and 670-fold, respectively), have a negligible effect on hGH-hGHR binding, and mildly retard the PRL-PRLR binding.

The differences in electrostatic contribution become obvious when the electrostatic surfaces of the subunits in the four systems are displayed. As shown in Fig. 2, *A–D*, the receptor binding sites on EPO and IL4 feature mostly positive electrostatic surfaces, which complement mostly negative electrostatic surfaces of the cytokine binding sites on EPOR and IL4Rα. This is a common feature of protein complexes with significant electrostatic rate enhancement (20). In contrast, the two sides of the interface in the hGH-hGHR and PRL-PRLR complexes lack such electrostatic complementarity, with both of the two electrostatic surfaces in each complex having mixed positive and negative regions (Fig. 2, *E–H*).

Charge complementarity index

As the results presented above demonstrate, electrostatic complementarity can provide significant enhancement of protein-protein association rates. We wanted to capture electrostatic complementarity without having to solve the Poisson-Boltzmann equation for the electrostatic surfaces. Here, we propose a charge complementarity index calculated by simply collecting the charge pairs across the binding interface within a distance cutoff. The CCI values for the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR complexes were 2.4, 0.5, –0.1, and –0.3, respectively, apparently correlating well ($R^2 = 0.89$) with the corresponding ΔG_{el}^* values (−4.2 kcal/mol, −2.4 kcal/mol, −0.1 kcal/mol, and 0.5 kcal/mol) (see Fig. S1 in the Supporting Material).

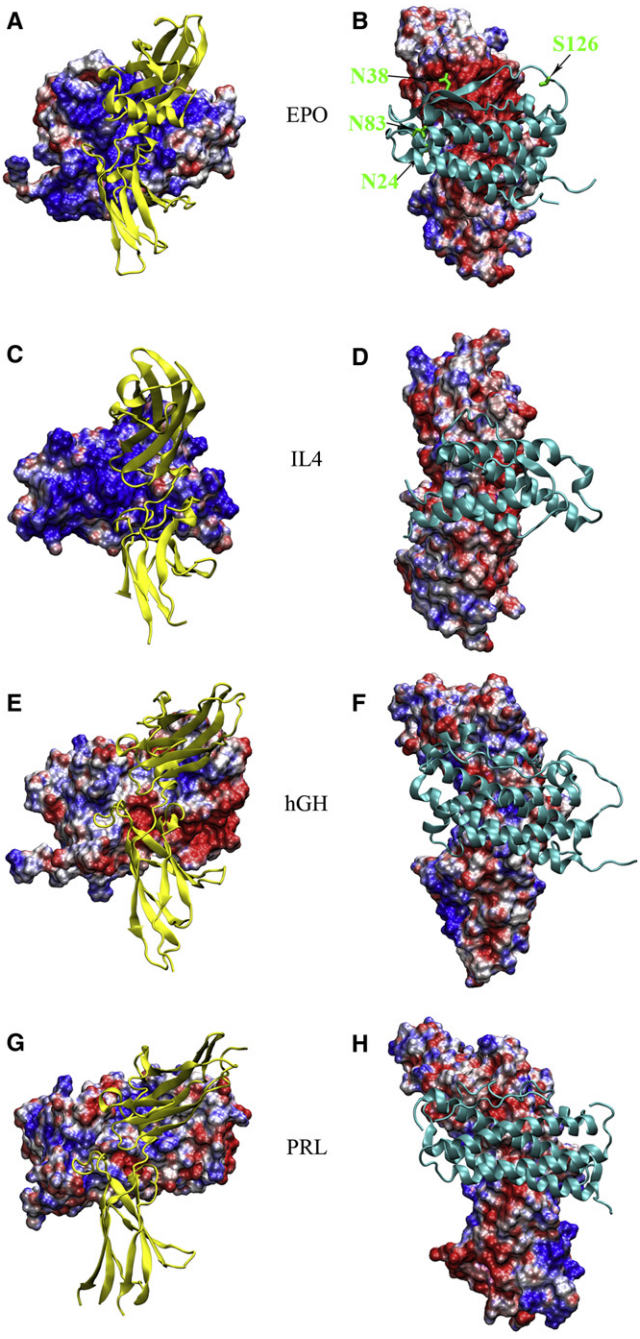

FIGURE 2 Electrostatic surfaces of the subunits in the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR complexes. (*A*) Electrostatic surface of EPO, with the bound EPOR shown as yellow ribbon. (*B*) Electrostatic surface of EPOR, with the bound EPO shown as cyan ribbon. The four glycosylation sites, N24, N38, N83, and S126, are shown by the side chains in green stick. (*C*) Electrostatic surface of IL4, with the bound IL4Rα shown as yellow ribbon. (*D*) Electrostatic surface of IL4Rα, with the bound IL4 shown as cyan ribbon. (*E*) Electrostatic surface of hGH, with the bound hGHR shown as yellow ribbon. (*F*) Electrostatic surface of hGHR, with the bound hGH shown as cyan ribbon. (*G*) Electrostatic surface of PRL, with the bound PRLR shown as yellow ribbon. (*H*) Electrostatic surface of PRLR, with the bound PRL shown as cyan ribbon. Electrostatic potential ranging from –5 k_BT/e to 5 k_BT/e is presented in a red/white/blue spectrum.

We also tested this CCI on a set of 100 other complexes with ΔG^*_{e1} values ranging from –7.1 to 3.4 kcal/mol. On this larger set CCI also showed reasonable correlation with ΔG^*_{e1} (data not shown).

Effects of ionic strength and charge mutations on EPO-EPOR binding rate

Experimentally it was observed that the EPO-EPOR association rate constant k_a reduced significantly, by 58-fold, when the ionic strength increased from 160 to 1010 mM, whereas the dissociation rate constant changed minimally (6). Qualitatively this is consistent with our previous finding (20,21,26,27) that the electrostatic enhancement of k_a decreases significantly with increasing ionic strength, as mobile ions screen the electrostatic interactions between the proteins. Fig. 3 shows that the measured ionic strength dependence of k_a is quantitatively reproduced well by our calculations, for both NGE and the NGE-Glu mutant.

At $I = 160$ mM, the calculated k_a of NGE-Glu is lower by 2.8-fold than the calculated k_a of NGE. This difference also compares favorably with the experimental counterpart, 3.4-fold (6). The four EPO glycosylation sites are located on the periphery of the EPOR binding site. Replacement by Glu residues at these sites will add negatively spots on the mostly positive electrostatic surface facing EPOR. Hence, the long-range electrostatic repulsion of these Glu residues by the negative electrostatic surface of EPOR accounts for the lower k_a of NGE-Glu. In terms of electric charges, the NGE-Glu mutant mimics a glycosylated EPO, and the EPOR binding rate constants of these two variants are also similar (6). Going from NGE-Glu to NGE corresponds to desialylation, which was found to increase EPO-EPOR binding affinity through increasing k_a (2,6). Our calculations on NGE-Glu and NGE here suggest that the increased k_a upon desialylation arises from removing

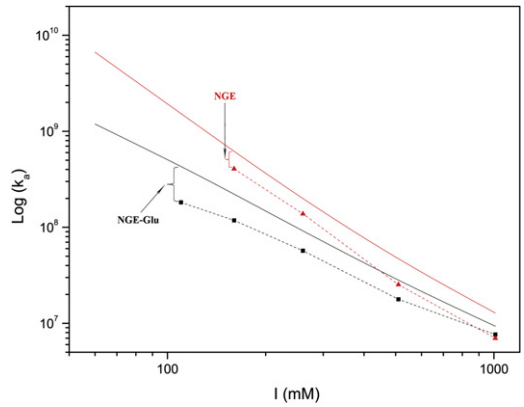

FIGURE 3 Comparison of ionic-strength dependences of calculated and experimental k_a for NGE and NGE-Glu binding to EPOR. The calculated k_a results are shown as continuous curves, whereas the experimental k_a results (6) are shown as symbols connected by dashes.

the negative charges on the sialyl groups, which otherwise would encounter repulsion from EPOR.

If instead of Glu the glycosylation sites are replaced by Lys, the positively charged residues are expected to enhance the attraction to EPOR through long-range electrostatic interactions. Our calculated k_a of NGE-Lys at I = 160 mM was higher by 2.5-fold than that of NGE. Curiously, the measured k_a values of NGE-Lys and NGE were essentially the same (6).

Effects of charge mutations on IL4-IL4Rα binding rate

In a previous study (20), we were able to reproduce experimental results for the effects of ionic strength and charge neutralizations on the IL4-IL4Rα association rate. These charge neutralizations were found to have modest effects (<1.5-fold reduction) in k_a. Here we studied the effects of charge reversals on IL4. The association rate constants of the IL4 K77E, R81E, K84D, and R85E mutants were calculated to be 3.9×10^6 $M^{-1}s^{-1}$, 6.9×10^5 $M^{-1}s^{-1}$, 2.9×10^6 $M^{-1}s^{-1}$, and 2.3×10^6 $M^{-1}s^{-1}$ (corresponding to reduction in k_a of 3.4-, 18.8-, 4.5-, and 5.7-fold), respectively. These results compare favorably with the experimental values, 4.4×10^6 $M^{-1}s^{-1}$, 3.2×10^6 $M^{-1}s^{-1}$, 2.4×10^6 $M^{-1}s^{-1}$, and 3.6×10^6 $M^{-1}s^{-1}$, respectively (8). The reduction in k_a is easily explained by the electrostatic surfaces. As shown in Fig. 2, *C–D*, the IL4Rα binding site on IL4 has a mostly positive electrostatic surface, which complements a mostly negative electrostatic surface of the IL4 binding site on IL4Rα. Therefore, replacement of the positively charged residues of IL4 by negatively charged residues change favorable electrostatic interactions into unfavorable ones.

Effects of charge mutations on hGH-hGHR binding rate

Effects on the hGH-hGHR binding rate by charge mutations at numerous positions on both hGH and hGHR were found to be minimal (less than twofold) (13–15). An affinity matured hGH variant was found to achieve affinity enhancement through a decrease in k_d, with little effect on k_a (15). Qualitatively, these observations are explained by our calculation result that electrostatic interactions contribute very little to the hGH-hGHR binding rate. Our calculations on these mutants confirmed that their k_a values differ from the counterpart of the wild-type complex by less than twofold.

A hGH-hGHR variant with putatively enhanced binding rate

Despite the inability of the large number of mutations tested so far (13–15) to produce significant rate enhancement, the results on the EPO-EPOR binding rate suggest that significant electrostatic rate enhancement can be introduced to hGH-hGHR binding through mutation. As noted previously, the main difference between the two systems is that EPO and EPOR have mostly positive and mostly negative electrostatic surfaces, respectively, across the binding interface, whereas the electrostatic surfaces of both hGH and hGHR have mixed positive and negative regions (Fig. 2, *A*, *B*, *E*, and *F*). We therefore designed charge mutations that would make one electrostatic surface (presumably that of hGH, to mimic EPO) mostly positive and the other electrostatic surface (presumably that of hGHR, to mimic EPOR) mostly negative. Inspection of the hGH electrostatic surface around the binding interface revealed that two negatively charged residues, D171 and E174, give rise to a negative region that is surrounded by a mostly positive periphery (Fig. 4, *A* and *B*). We reversed the charges of these residues by mutating them to Arg. Similarly, on the hGHR electrostatic surface, a positive region due to R43 and R217 is surrounded by a mostly negative periphery (Fig. 4, *A* and *C*). We reversed these charges by mutating them to Glu. With the charge reversal on these four residues in the hGH-hGHR interface, the two proteins now have both strong local electrostatic interactions (Fig. 4 *D*) and good complementary electrostatic surfaces (Fig. 4, *E* and *F*).

The designed hGH-hGHR mutant was found to have significant attraction. The electrostatic interaction free energy in the transient complex, ΔG^*_{e1}, changed from -0.1 kcal/mol to -2.7 kcal/mol at $I = 160$ mM. As a result, the calculated k_a changed from 2.0×10^5 $M^{-1}s^{-1}$, and 1.6×10^7 $M^{-1}s^{-1}$, an 83-fold increase. As expected, the designed mutations also improved CCI, from –0.1 to 0.8, suggesting that CCI might serve as a guide for designing mutants with enhanced association rates.

DISCUSSION

Protein-protein association is at the center of diverse biological processes ranging from enzyme catalysis/inhibition to regulation of immune response by cytokines. The association rates often play a critical role in such processes. Therefore, theoretical prediction of the association rate constants is of great importance (31). A widely used approach for calculating k_a is by simulating the translational and rotational Brownian motion of the subunits (32–35). This approach often involves adjusting parameters for specifying the conditions for association to achieve optimal agreement with experimental results, thereby compromising the predictive power. Moreover, this approach is computationally expensive (e.g., requiring ~11 weeks running on 10 8-core Intel CPUs in a recent study (35)). Our approach based on the transient-complex theory overcomes both of these obstacles (19–22). Most importantly, the approach allows us to tease out the contributions to k_a, thus providing insight into the control of association rate constants. This ability

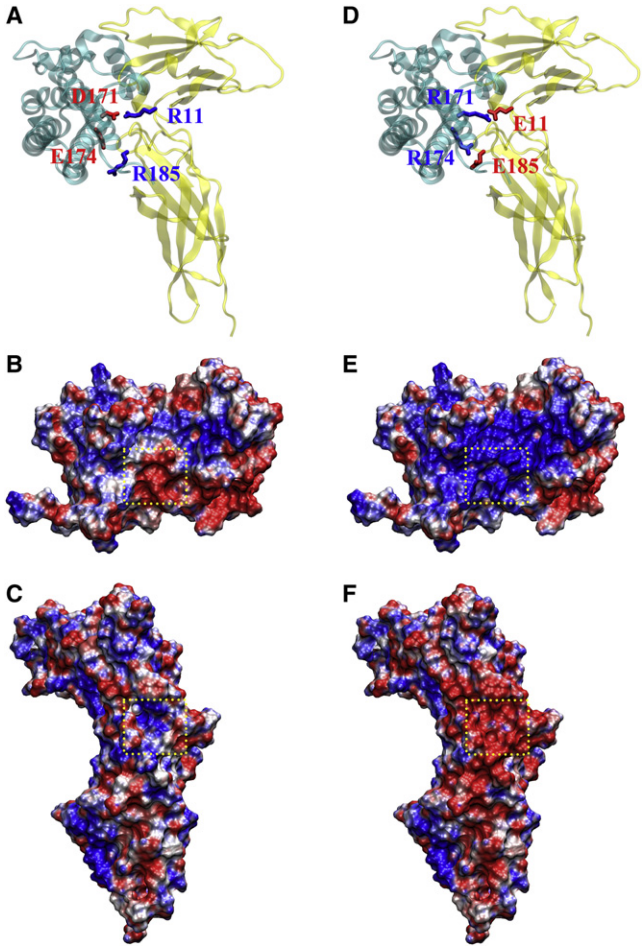

FIGURE 4 Electrostatic surfaces of the wild-type hGH-hGHR complex and the designed mutant. (*A*) The location of the four residues to be mutated in the native complex. (*B*) The negative region, defined by a box with dashed borders, on the hGH electrostatic surface around D171 and E174. (*C*) The positive region, defined by a box with dashed borders, on the hGHR electrostatic surface around R43 and R217. (*D*) The local interactions of the mutated residues in the native complex. (*E*) The electrostatic surface of the mutated hGH. Note the contrast of the boxed region here and in *B*. (*F*) The electrostatic surface of the mutated hGHR. Note the contrast of the boxed region here and in *C*. Electrostatic potential ranging from –5 k_BT/e to 5 k_BT/e is presented in a red/white/blue spectrum.

is well illustrated here by our study of the EPO-EPOR, IL4-IL4Rα, hGH-hGHR, and PRL-PRLR systems. We have not only quantitatively reproduced the observed 5000-fold differences in rate constants without adjusting any parameters, but also provided a physical explanation for the vast rate differences.

Is there a physiological reason for the vast differences in the receptor-binding rate constants of EPO, IL4, hGH, and PRL? Rapid association of some proteins (e.g., barnase and bastar) has been suggested to play a critical physiological role (e.g., for self-defense) (31). There is no evidence indicating that this is the case for the cytokine-receptor systems studied here. On the other hand, the clearance of receptors due to degradation or other reasons presents another limiting factor. That is, the receptors are useful only if cytokine binding occurs before their clearance. The half-life $\tau_{1/2}$ of the receptor sets a lower limit, $k_{1/2} \equiv 1/\tau_{1/2}$ for the pseudo-first order receptor-binding rate constant k_aC_{cyt}, where C_{cyt} is the cytokine circulation concentration. One expects $k_aC_{cyt} \geq k_{1/2}$. We now check this expectation on the four cytokine-receptor systems studied here.

EPO presents in the plasma at very low concentrations, ~0.8–4 pM (36), even though it stimulates the very fast production of ~2.3 million red cells/s. With observed $k_a = 4.0 \times 10^8\ M^{-1}s^{-1}$ (6), we find $k_aC_{cyt} = 3.2 \times 10^{-4}\ s^{-1}$. The half-time of EPOR is ~1.5 h (37), correspondingly $k_{1/2} = 1.9 \times 10^{-4}\ s^{-1}$. It thus seems that the pseudo-first order rate constant k_aC_{cyt} is barely enough to pass the lower limit set by the half-life of EPOR. To make this happen, the high k_a is required to accommodate the low C_{cyt}. The results for the IL4- IL4Rα system are similar. The minimum IL4 concentration required of T-cell proliferation is ~20 pM (38). With $k_a = 1.3 \times 10^7\ M^{-1}s^{-1}$ (8), we find $k_aC_{cyt} = 2.6 \times 10^{-4}\ s^{-1}$. This is comparable to the limit $k_{1/2} = 0.8 \times 10^{-4}\ s^{-1}$ set by the 3.5-h half-life of IL4Rα (39). In contrast, the two slow binding cytokines are present at much higher concentrations, both at ~10 nM (40,41). For hGH-hGHR binding, with $k_a = 3.2 \times 10^5\ M^{-1}s^{-1}$ (15) we find $k_aC_{cyt} = 3.2 \times 10^{-3}\ s^{-1}$, to be compared with the limit $k_{1/2} = 0.6 \times 10^{-3}\ s^{-1}$ set by the 30-min half-life of hGHR (42). For PRL-PRLR binding, with $k_a = 8.0 \times 10^4\ M^{-1}s^{-1}$ (17) we find $k_aC_{cyt} = 0.8 \times 10^{-3}\ s^{-1}$, to be compared with the limit $k_{1/2} = 0.4 \times 10^{-3}\ s^{-1}$ set by the 40-min half-life of PRLR (42). We thus observe that the 5000-fold differences in k_a of the four cytokine-receptor systems are inversely correlated with a 10,000-fold variation in C_{cyt}. We further conclude that the receptor-binding rate constants and the circulation concentrations of the different cytokines are evolutionarily tuned to ensure that all the receptors produced would participate in cytokine binding rather than being wasted.

The design of mutants with higher association rates may be important for the increased rates alone. Such designed mutants have the additional advantage that the binding affinity is also enhanced (assuming that the dissociation rate constant is not adversely affected). Extensive mutational studies on hGH and hGHR (13–15) have not produced any mutant with a significantly enhanced k_a but have produced a mutant with enhanced binding affinity through slowing down dissociation (15). Here, we designed an hGH-hGHR mutant that is predicted to have nearly 100-fold enhancement in k_a. Our design strategy follows that of Schreiber and co-workers (43,44), by focusing on charge mutations around the interface. In the systems studied by Schreiber and co-workers, the binding site on the first subunit is dominated by one type of charge, whereas the binding site on the second subunit contains mixed positive and negative charges. Therefore, they focused on the

latter binding site to introduce charges complementary to those on the first subunit. In the hGH-hGHR complex studied here, both binding sites contain mixed positive and negative charges, explaining why previous mutational studies failed to produce mutants with enhanced k_a. Nevertheless, we were able to design double mutations on the two subunits to produce two electrostatic surfaces that are mostly positive and mostly negative, respectively. Our design approach is expected to be applicable to many other systems.

It has been suggested that rapid association is as important as high affinity in the proper functioning of proteins (31). Manipulating association rate constants of various components thus presents unique opportunities for the control of protein functions. The predictive power of our transient-complex theory for calculating protein-protein association rate constants have been demonstrated in previous studies (20,21,27) and further demonstrated here by our results on four cytokine-receptor systems. We also designed a mutant of the hGH-hGHR complex with a putative 83-fold increase in association rate. To guide such design, we have proposed a simple charge complementarity index, based on the charge distribution around the binding interface.

On the four cytokine-receptor systems themselves, many questions remain. For example, the mechanisms by which the binding of these cytokines to the extracellular domains of their receptors transmits signals through cell membranes are still unknown. The binding process studied here, forming the high-affinity 1:1 complex, is only part of the mechanisms. It is known that a second receptor molecule subsequently binds to the low-affinity second site of the cytokine in the 1:1 complex, and it is the resulting dimerization of the receptor molecules that initiates signaling (11). Blocking the binding of the second receptor is a focus of cytokine antagonist design (14,17,45–47). We plan to study these downstream steps in the future.

In conclusion, we have applied the transient-complex theory to quantitatively rationalize the 5000-fold differences in receptor binding rate constants among four cytokines and have provided a physical explanation for the vast differences. The EPO-EPOR and IL4-IL4Rα complexes have a mostly positive electrostatic surface on one side of the interface and a mostly negative electrostatic surface on the other, a feature common to protein complexes with significant electrostatic rate enhancement. In contrast, the electrostatic surfaces on both sides of the interface in the hGH-hGHR and PRL-PRLR complexes have mixed positive and negative regions. We have uncovered that the vast differences in k_a are anticorrelated with the equally vast differences in cytokine circulation concentration and conclude that both are evolutionarily tuned. The vast differences in k_a and other results presented here suggest that protein charges can be manipulated to tune k_a and control biological function.

SUPPORTING MATERIAL

A figure is available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(11)00787-9.

This work was supported in part by National Institutes of Health grant GM58187.

REFERENCES

1. Dubé, S., J. W. Fisher, and J. S. Powell. 1988. Glycosylation at specific sites of erythropoietin is essential for biosynthesis, secretion, and biological function. *J. Biol. Chem.* 263:17516–17521.
2. Tsuda, E., G. Kawanishi, …, R. Sasaki. 1990. The role of carbohydrate in recombinant human erythropoietin. *Eur. J. Biochem.* 188:405–411.
3. Philo, J. S., K. H. Aoki, …, J. Wen. 1996. Dimerization of the extracellular domain of the erythropoietin (EPO) receptor by EPO: one high-affinity and one low-affinity interaction. *Biochemistry.* 35:1681–1691.
4. Syed, R. S., S. W. Reid, …, R. M. Stroud. 1998. Efficiency of signaling through cytokine receptors depends critically on receptor orientation. *Nature.* 395:511–516.
5. Middleton, S. A., F. P. Barbone, …, L. K. Jolliffe. 1999. Shared and unique determinants of the erythropoietin (EPO) receptor are important for binding EPO and EPO mimetic peptide. *J. Biol. Chem.* 274:14163–14169.
6. Darling, R. J., U. Kuchibhotla, …, J. M. Beals. 2002. Glycosylation of erythropoietin affects receptor binding kinetics: role of electrostatic interactions. *Biochemistry.* 41:14524–14531.
7. Frank, S. J. 2002. Receptor dimerization in GH and erythropoietin action—it takes two to tango, but how? *Endocrinology.* 143:2–10.
8. Wang, Y. H., B. J. Shen, and W. Sebald. 1997. A mixed-charge pair in human interleukin 4 dominates high-affinity interaction with the receptor alpha chain. *Proc. Natl. Acad. Sci. USA.* 94:1657–1662.
9. Hage, T., W. Sebald, and P. Reinemer. 1999. Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface. *Cell.* 97:271–281.
10. de Vos, A. M., M. Ultsch, and A. A. Kossiakoff. 1992. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science.* 255:306–312.
11. Fuh, G., B. C. Cunningham, …, J. A. Wells. 1992. Rational design of potent antagonists to the human growth hormone receptor. *Science.* 256:1677–1680.
12. Clackson, T., and J. A. Wells. 1995. A hot spot of binding energy in a hormone-receptor interface. *Science.* 267:383–386.
13. Pearce, Jr., K. H., M. H. Ultsch, …, J. A. Wells. 1996. Structural and mutational analysis of affinity-inert contact residues at the growth hormone-receptor interface. *Biochemistry.* 35:10300–10307.
14. Clackson, T., M. H. Ultsch, …, A. M. de Vos. 1998. Structural and functional analysis of the 1:1 growth hormone: receptor complex reveals the molecular basis for receptor affinity. *J. Mol. Biol.* 277:1111–1128.
15. Bernat, B., M. Sun, …, A. A. Kossiakoff. 2004. Dissecting the binding energy epitope of a high-affinity variant of human growth hormone: cooperative and additive effects from combining mutations from independently selected phage display mutagenesis libraries. *Biochemistry.* 43:6076–6084.
16. Walsh, S. T. R., J. E. Sylvester, and A. A. Kossiakoff. 2004. The high- and low-affinity receptor binding sites of growth hormone are allosterically coupled. *Proc. Natl. Acad. Sci. USA.* 101:17078–17083.
17. Jomain, J. B., E. Tallet, …, V. Goffin. 2007. Structural and thermodynamic bases for the design of pure prolactin receptor antagonists: x-ray structure of Del1-9-G129R-hPRL. *J. Biol. Chem.* 282:33118–33131.
18. van Agthoven, J., C. Zhang, …, I. Broutin. 2010. Structural characterization of the stem-stem dimerization interface between prolactin

receptor chains complexed with the natural hormone. *J. Mol. Biol.* 404:112–126.

19. Alsallaq, R., and H. X. Zhou. 2007. Energy landscape and transition state of protein-protein association. *Biophys. J.* 92:1486–1502.
20. Alsallaq, R., and H.-X. Zhou. 2008. Electrostatic rate enhancement and transient complex of protein-protein association. *Proteins.* 71:320–335.
21. Qin, S. B., and H. X. Zhou. 2008. Prediction of salt and mutational effects on the association rate of U1A protein and U1 small nuclear RNA stem/loop II. *J. Phys. Chem. B.* 112:5955–5960.
22. Zhou, H. X. 2010. Rate theories for biologists. *Q. Rev. Biophys.* 43:219–293.
23. McCoy, A. J., V. Chandana Epa, and P. M. Colman. 1997. Electrostatic complementarity at protein/protein interfaces. *J. Mol. Biol.* 268: 570–584.
24. Fiser, A., R. K. G. Do, and A. Sali. 2000. Modeling of loops in protein structures. *Protein Sci.* 9:1753–1773.
25. Zhou, X. Z. 1995. Calculation of translational friction and intrinsic viscosity. II. Application to globular proteins. *Biophys. J.* 69:2298–2303.
26. Alsallaq, R., and H. X. Zhou. 2007. Prediction of protein-protein association rates from a transition-state theory. *Structure.* 15:215–224.
27. Qin, S. B., and H. X. Zhou. 2009. Dissection of the high rate constant for the binding of a ribotoxin to the ribosome. *Proc. Natl. Acad. Sci. USA.* 106:6974–6979.
28. Baker, N. A., D. Sept, …, J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA.* 98:10037–10041.
29. Cornell, W. D., P. Cieplak, …, P. A. Kollman. 1995. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* 117:5179–5197.
30. Bondi, A. 1964. van der Waals volumes and radii. *J Phys Chem.* 68:441–451.
31. Schreiber, G., G. Haran, and H. X. Zhou. 2009. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* 109:839–860.
32. Gabdoulline, R. R., and R. C. Wade. 1997. Simulation of the diffusional association of barnase and barstar. *Biophys. J.* 72:1917–1929.
33. Elcock, A. H., R. R. Gabdoulline, …, J. A. McCammon. 1999. Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin. *J. Mol. Biol.* 291:149–162.
34. Gabdoulline, R. R., and R. C. Wade. 2001. Protein-protein association: investigation of factors influencing association rates by Brownian dynamics simulations. *J. Mol. Biol.* 306:1139–1155.
35. Frembgen-Kesner, T., and A. H. Elcock. 2010. Absolute protein-protein association rate constants from flexible, coarse-grained Brownian dynamics simulations: the role of intermolecular hydrodynamic interactions in barnase-barstar association. *Biophys. J.* 99:L75–L77.
36. Lappin, T. R., A. P. Maxwell, and P. G. Johnston. 2002. EPO's alter ego: erythropoietin has multiple actions. *Stem Cells.* 20:485–492.
37. Wickrema, A., S. B. Krantz, …, M. C. Bondurant. 1992. Differentiation and erythropoietin receptor gene expression in human erythroid progenitor cells. *Blood.* 80:1940–1949.
38. Letzelter, F., Y. H. Wang, and W. Sebald. 1998. The interleukin-4 site-2 epitope determining binding of the common receptor gamma chain. *Eur. J. Biochem.* 257:11–20.
39. Husain, S. R., P. Leland, …, R. K. Puri. 1996. Transcriptional up-regulation of interleukin 4 receptors by human immunodeficiency virus type 1 tat gene. *AIDS Res. Hum. Retroviruses.* 12:1349–1359.
40. Ehrlich, R. M., and P. J. Randle. 1961. Immunoassay of growth hormone in human serum. *Lancet.* 2:230–233.
41. Villalba, M., M. T. Zabala, …, J. P. Garcia-Ruiz. 1991. Prolactin increases cytosolic free calcium concentration in hepatocytes of lactating rats. *Endocrinology.* 129:2857–2861.
42. Baxter, R. C. 1985. Measurement of growth hormone and prolactin receptor turnover in rat liver. *Endocrinology.* 117:650–655.
43. Selzer, T., S. Albeck, and G. Schreiber. 2000. Rational design of faster associating and tighter binding protein complexes. *Nat. Struct. Biol.* 7:537–541.
44. Kiel, C., T. Selzer, …, C. Herrmann. 2004. Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proc. Natl. Acad. Sci. USA.* 101:9223–9228.
45. Wenzel, S., D. Wilbraham, …, M. Longphre. 2007. Effect of an interleukin-4 variant on late phase asthmatic response to allergen challenge in asthmatic patients: results of two phase 2a studies. *Lancet.* 370:1422–1431.
46. Tony, H. P., B. J. Shen, …, W. Sebald. 1994. Design of human interleukin-4 antagonists inhibiting interleukin-4-dependent and interleukin-13-dependent responses in T-cells and B-cells with high efficiency. *Eur. J. Biochem.* 225:659–665.
47. Andrews, A. L., J. W. Holloway, …, D. E. Davies. 2006. IL-4 receptor alpha is an important modulator of IL-4 and IL-13 receptor binding: implications for the development of therapeutic targets. *J. Immunol.* 176:7456–7461.