PROTEIN-PROTEIN RECOGNITION:

The Neonatal Fc Receptor and Immunoglobulin G

Thesis by

Warham Lance Martin

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Cailfornia Institute of Technology

Pasadena, California

2001

(Submitted May 17, 2001)

To my Father,

To my Mother,

and Barbara

It is not the critic who counts, not the man who points out how the strong man stumbles or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood, who strives valiantly, who errs and comes short again and again because there is no effort without error and short comings; but who actually strives to do the deed, who knows the great devotion, who spends himself in a worthy cause, who at best knows in the end the high achievement of triumph and who at worst, if he fails while daring greatly, knows his place shall never be with those timid cold souls who know neither victory nor defeat.

<div align="center">Theodore Roosevelt</div>

These good words were given to me by Jay Hoffman, a dear friend, for Christmas in 1996. For better or worse they have been a lodestone during my thesis.

Acknowledgments

I have looked forward to writing this section of the thesis for a long time. No, not because it was the last part to be written, how cynical. What can I say, I'm gonna groove on thanking ya'll. So I put Led Zepplin II in the CD player (Thanks Barb!) and the head phones on my head and off we go.

First, I offer a general thanks to the Caltech community. I grew up at 1056 N. Holliston, here in Puh-sah-di-nah. To grow up in the shadow of the school of Feynman, Pauling, Delbruck, and all the rest, well it made the idea of coming here seem rather unattainable. But that awe was nothing compared to what I found, not a thing. This should not be hard to say but it is, so many people here shown me how to become better. Some people here inspire with brilliance, others with quiet competence, some with depth of understanding, others with the capacity to care, and some with ability to smile through adversity. Even and perhaps especially during times when I felt like a black sheep, the people of Caltech have reached me, made me feel like I belonged, and compelled me to try to improve. As a child of the Book I grew up looking to the hills for strength, now I look to the people, the people of Caltech, for that precious gift. For the strength you gave, for the audacious talents you exhibited, and for the indomitable love I have felt while here I thank you.

When you were a kid, did you like opera? Me neither, all that warbling, good grief. Fortunately for me, I have sisters who sing, a lot. I learned from them how difficult it is to sing with volume and broad range in pitch for long periods of time. They also taught me the stories of opera, stories of human folly and heroism, pagentry and sublime emotion. Then they described the inside stories, the tragic lives of the

composers, the romance and intrigue involved in operatic production, and the music, what it means to those who sing it. If you know these things, can you listen to the end of the Mozart's Marriage of Figaro, where the philandering Count begs his wife's forgiveness and she bestows it, without weeping? The human frailty, the human regret, the human magnificence, distilled by genius into an aria for the ages. It is hard for an only boy to admit that his sisters gave him anything but how can I deny this gift?

I mention how I learned about opera to explain what I have learned from Pamela, my adviser. It is easy to praise her, her brilliance, her competence, her achievements, everyone does. I want to tell you what it is like for someone on the inside, in the lab. When I started here I saw science as a means to an end, a way to improve the world. I learned from Pamela, a deeper truth. Not solely the merit of pursuing science for its own sake, I learned from Pamela the necessity of it. Science, good science, is hard. To push back the borders of human understanding, even the slightest bit, you must love your science with ferocity unceasing. This I learned from mi Maestra, knowing the dedication required to make science happen, you should weep with joy at progress no matter how meager, thank you Boss for this. If any of the rest of us could muster her commitment (I know, I try and fail daily) we would slay the dragons of disease and hunger with speed and grace.

It is probably not difficult to imagine how the lab of such an adviser would rock, well it does. It is a mag/lev train of progress, a support group for frustration, a deep well of experience, a place for fellowship, and it has been for me a surrogate family. Con Vaughn's arguments, Jose's laugh or his Witch project, Tara's jokes and spiders, Watston's goofy stories, Yangster's cowboy hat, Zsuzsi's rants and venting, and Tony's

question about sequence and structure without having him give you the answer, first. I'm telling you it can't be done. More than a cohabitant, David has been a cherished sounding board for all manner of ruminations on my part. What a pleasure to have such a filter for the froth from one's mind, thank you Master Mathogian. Thanks, Ben (Teflon) Bornstein for the lifting, the understanding, the late night C jam sessions, and the PCA, as to the teflon, this too shall pass, in fact it already has, dig it. I thank Dr. Mary Dickinson for her preternatural skill with the two-photon microscope. Hey Scott, I know you know your people got game, but Mary D has intangibles in spades. Doing the FRET experiments with her is an unexpected pleasure. I thank Pavel Strop and Dan Bolon for many helpful discussions, references, Dan for the Chupa Fangos, and Pavel for recent calculations. I thank Dr. Shelley Diamond, for all her help over the years with FACS analysis. Ellen, I know you know that Shelley is the bomb, did you know she moonlights as a Mother Teresa for wayward graduate students, thank you so much for sharing her with the rest of us.

That brings me to my committee. Ya'll don't know this but I choose you not solely for your science, I actually placed a greater premium on whether I felt in my gut that I could trust you. Well, as has been typical of my experience here, my penurious estimates of people have been washed away in the flood of their Brobdingnagian actual worth. Doug, I hope that you don't have a TA like me for a long time, I know that wish will make your life a little easier. Thanks for your patience during my extravagant flights of fancy, I am learning, albeit at a glacial pace the merits of the intellectual discipline that you (and Pamela, for that matter) exercise so well. You may worry about filling Linus' shoes, don't, you've got it covered, hombre. Your students, your lab, our lab, well, we

say it like this: When you talk to Doug, you go in knowing he's kind and that he'll be helpful, but then you're there and you're swimming in this Marianas Trench of structural understanding (I quote James Brown, "Good God Ya'll"). Ellen, dear Ellen, did you know that your words, your kind words, have kept me here in this game more than once, thank you. Your encyclopedic knowledge of the byzantine thicket of quicksand that is immunology is not outdone and your willingness to share it is an underappreciated gift to all of us at Caltech. Scott, going into your lab is like stumbling through some cosmic rabbit hole and being confronted not by a smiling cat from Cheshire but by instruments and people of mythic capabilities. Thanks for the enthusiasm about my crazy FRET idea, and for the hazing, the idiosyncracies of my psyche are such that the verbal hazing is as cherished as a warm hearth.

My penultimate gratitude I extend to those who, in addition to those above have helped me during the tough times. As Anthony says, I have a broader range of moods than most and well, the lows are pretty grim. To Randal Bass, for your witness, the philos, the waves, for your soul of titanium, thanks and pax dei vobiscum, Cuddles. Jeroen Corver, my man from Rotterdam, for the hikes, the horseback riding, the surfing and the example of your happy, intelligent, quietude, thank you. To Carol Povenmire and Janis Schonauer, for listening and hearing through my pain, fear, and anger, thank you both. I cannot repay you but by making your help unnecessary. To Dean, Sally, and Peg for your warm reception, your inclusion in your family, you were a lighthouse in a bitter Nor'easter, thank you. To Yoda, Kut, and Jay, for love across decades that I wear with apologies to the people of Ephesis, like a shield of faith, thank you. I'll see you in Seattle, 'round the fire, mon frere. Finally, thanks to my Dad and Mom who by example

revealed that if you work hard, try to do right, and don't get consumed by the "pride that grows on the heart of a man like lard on the heart of a pig,"(thanks Solzhenitsyn) that all, even those who hold you in contempt, will see the truth of you (though it may take fifty years).

Now for the best, who is the maelstrom, who is the calm eye of the hurricane of experience that has been my thesis? Let me tell the story like this, sometimes you just get plain lucky. My Mom had four miscarriages before she gave birth to five children. Her second, my brother Travis, died before six months due to a malformation of his heart. At five weeks, I could not eat for spitting up, I had surgery to remove a plyoric stenosis. When I was a kid, with three sisters I wanted nothing more than a brother, but my parents were quite and well done generating progeny. At the age of seven my wish came true, I found not one but two brothers down the street (Yoda and Kut above). No one deserves to be so lucky twice in one lifetime, and yet here I stand, luckier than twice that by a bushel and a peck. You may have noticed the number of people I thanked above for listening, and the length of these acknowledgements is evidence of a certain loquacious tendency on my part. Imagine being my poor significant other, the woman has listened to more than any legion ought to have. To Barbara, for listening, loving, talking, teaching, surfing, skiing, Bayesian reasoning, AI, walking, planning, waiting, hoping, sharing, for Hawaii, New York, the Malleck farm, for the schmutz, for Boggle, Scrabble, Bridge, Monopoly, for dancing and pool at the Mix, for all that I have forgotten, for the way your eyes still sparkle at me no matter how old my patter gets, thank you. I know I owe you, but unlike the rest, I may have the pleasure of repaying you up at Berkeley. Long may we laugh and love.

To anyone else who has made it safe thus far, thanks for your effort. If you intend to venture further, I borrow from good Prince Hal and urge you to go, "once more into the breach." Bona ventura.

Abstract

The neonatal Fc receptor (FcRn) binds the Fc portion of immunoglobulin G (IgG) at the acidic pH of endosomes or the gut and releases IgG at the alkaline pH of blood. FcRn is responsible for the maternofetal transfer of IgG and for rescuing endocytosed IgG from a default degradative pathway. We investigated how FcRn interacts with IgG by constructing a heterodimeric form of the Fc (hdFc) that contains one FcRn binding site. This molecule was used to characterize the interaction between one FcRn molecule and one Fc and to determine under what conditions FcRn forms a dimer. The hdFc binds one FcRn molecule at pH 6.0 with a K_d of 80 nM. In solution and with FcRn anchored to solid supports, the heterodimeric Fc does not induce a dimer of FcRn molecules. FcRn-hdFc complex crystals were obtained and the complex structure was solved to 2.8 Å resolution. Analysis of this structure refined the understanding of the mechanism of the pH-dependent binding, shed light on the role played by carbohydrates in the Fc binding, and provided insights on how to design therapeutic IgG antibodies with longer serum half-lives. The FcRn-hdFc complex in the crystal did not contain the FcRn dimer. To characterize the tendency of FcRn to form a dimer in a membrane we analyzed the tendency of the hdFc to induce cross-phosphorylation of FcRn-tyrosine kinase chimeras. We also constructed FcRn-cyan and FcRn-yellow fluorescent proteins and have analyzed the tendency of these molecules to exhibit fluorescence resonance energy transfer. As of now, neither of these analyses have lead to conclusive results. In the process of acquiring the context to appreciate the structure of the FcRn-hdFc interface, we developed a study of 171 other nonobligate protein-protein interfaces that includes an original principal component analysis of the quantifiable aspects of these interfaces.

Table of Contents

List of Figures and Tables

Chapter 6

Appendix A

Chapter 1:

Introduction

Protein-Protein Recognition: Areview of the analyses of collections of crystal structures

It is difficult to overstate the importance of the formation of protein-protein complexes in biological processes. To determine whether a cell in the body has been virally infected, a T-cell receptor on a cytotoxic T-cell must bind a major histocompatibility complex (MHC) class I molecule on the infected cell (Garboczi et al., 1998). In order to couple ATP hydrolysis with the electron transfer necessary to fix atmospheric nitrogen, the nitrogenase iron protein must bind the nitrogenase molybdenum-iron protein (Schindelin et al., 1997). In order to activate T-helper cells, a central step in adaptive immune responses, the nuclear factor of activated T-cells, NFAT, must bind the complex of Fos and Jun proteins to activate the expression of many immune response genes (Chen et al., 1998). To rearrange the actin cytoskeleton, necessary for cellular motility, the small G-proteins of the Rho family, RhoA and RhoGAP form a complex and instigate a phosphorylation cascade by enhancing the rate of GTP hydrolysis (Rittinger et al., 1997). To rescue serum immunoglobulin gamma, (IgG) from a default degradation pathway, enhancing the role of these antibodies in fighting disease, the neonatal Fc receptor (FcRn) binds IgG in cellular endosomes and releases IgG at the cell surface (Martin et al., 2001). The formation of protein-protein complexes occurs in every organelle, every cell, every tissue, and every organism. Understanding how these complexes form and determining their structures is sufficently complicated that the complexes are usually studied one at a time, as in this thesis. However, some aspects of protein-protein complexes are studied collectively.

The current understanding of protein-protein recognition stems in part from the analyses of collections of crystal structures (Stites, 1997). To provide context for understanding the interaction between FcRn and IgG, the last chapter of this thesis is such an analysis. This introduction reviews related analyses of protein crystal structures. This introduction is an effort to show how analyzing collections of crystal structures has improved our understanding of how proteins are folded, oligomer subunits associate, protein structure is related to function, different groups of proteins may be distinguished, and what different aspects of a protein may be assessed quantitatively.

Multiple crystal structure analyses: Learning how proteins fold

Analyses of protein crystal structures have been undertaken to compare the data with existing theories on how proteins are folded. Before the existence of large amounts of protein structure data, studies of protein structure and theories of protein folding emphasized different aspects of a protein's fold. Pauling anticipated the importance of hydrogen bonding to secondary structure by predicting the formation of α-helices and β-sheets in polypeptide chains (Pauling et al., 1951; Pauling and Corey, 1951). Kauzmann predicted there would be entropic contributions to folding as proteins folded to shield large amounts of their hydrophobic surface area from the aqueous solvent (Kauzmann, 1959). Examination of the structure of myoglobin reported in 1960 appeared to confirm the relevance of both hypotheses, the structure is dominated by α-helices, while the nonpolar aliphatic and aromatic residues were primarily buried in the interior while the polar residues were exposed to solvent (Perutz et al., 1965).

To quantitatively verify this qualitative observation of hydrophobic burial made by Perutz, an algorithm was developed to measure the amount of protein surface

accessible to solvent on an atom-by-atom basis (Lee and Richards, 1971). This method generates a solvent accessible surface by building spheres centered on each crystal structure coordinate with radii equal to the sum of the van der Waal radii of the chemical group centered on that atom and the radius of water. This atom-by-atom analysis classed the protein atoms observed in crystal structures into three groups: all carbon atoms were considered nonpolar, all oxygen, nitrogen, and sulfur atoms were considered polar, except the carboxylate oxygens and the amino and guanadinium nitrogen atoms that were considered charged. With this algorithm Lee and Richards examined the structures of myoglobin, lysozyme, and ribonuclease and found that a large fraction of the surface was nonpolar and a large fraction of the interior surface was polar (Lee and Richards, 1971). This quantitative review of a qualitative observation produced a somewhat unexpected result.

Resolving this apparent paradox of the preference of nonpolar residues for the protein interior and the presence in the interior of proteins of large amounts of polar surface area required a refinement in the way we describe how proteins fold (Chothia, 1976). Chothia observed that the formation of secondary structure buries large amounts of polar surface area. This burial of polar surface area is energetically offset by the formation of a large number of hydrogen bonds, and the formation of the hydrogen bonds diminishes the polar, hydrophilic, character of the surfaces involved. In contrast, the formation of tertiary structure buries almost exclusively nonpolar surface area. This hydrophobic surface area buried is largely associated with amino acid side chains, and is driven by the hydrophobic effect.

Much as the measuring the surface area was understanding the hydrophobic effect measuring the volumes occupied by atoms, or how well the atoms are packed, was helpful to understanding the importance of the van der Waals forces. Computational efforts to assess packing have employed Voronoi polyhedra to measure the volumes of atoms (Chothia, 1975; Harpaz et al., 1994). A Voronoi polyhedron is generated around an atom in a protein in two steps. First, lines from that atom to each of its neighbors are drawn and second planes normal to these lines are placed according to the van der Waals radii of the two atoms. Each plane extends until it intersects another. This method is useful for judging atom density because, except for a vertex error due to differences in atom radii, it includes all the space around the atoms in the calculation (Gerstein et al., 1995).

Previous investigations into the volumes of amino acids in proteins (Chothia, 1975) were extended more recently in an analysis of 108 high resolution (1.0 Å – 1.9 Å) crystal structures (Harpaz et al., 1994). This study found that hydrophobic residues undergo a reduction in volume as a protein folds, which is somewhat offset by the increase in the volume of polar and charged residues. This is consistent with the small overall change in protein volume that occurs when proteins fold. Further calculations in this study show that amino acids in protein interiors occupy 4% less volume than in amino acid crystals. When compared with the 15% reduction in volume exhibited when organic solvents solidify, these results suggest that organic solvent molecules are significantly less well packed than a protein interior. This suggests that approximating protein folding as the transfer of a polypeptide chain from aquaeous buffer to organic solvent may under estimate the importance of van der Waals forces in protein folding.

These results support the proposition that protein stability is more accurately modeled by dissolving crystalline cyclic dipeptides in water (Murphy and Gill, 1991).

Multiple crystal structure analyses: Learning how oligomer subunits associate

The preference of nonpolar molecules for a nonaquaeous environment predicted by Kauzmann to be important for protein folding was also proposed as the driving force in protein-protein recognition (Janin and Chothia 1975). If this hydrophobic effect was to be considered the predominant source of energy for the interactions between polypeptide chains, then their association would less likely to be dominated by the formation of extensive hydrogen bond as seen in secondary structure or salt bridge networks. A study of 23 oligomers counted the number of hydrogen bonds formed between the polypeptide chains and found between none and 1 hydrogen bond per 200 Å^2 and 78% of these hydrogen bonds were formed by interaction between side-chains (Janin et al., 1988). This dearth of hydrogen bonds, especially those derived from contact between main-chain atoms, was consistent in oligomeric interfaces ranging from 700 Å^2 to 10,000 Å^2 in size. This confirmed the importance of the hydrophobic effect for the interactions between the polypeptide chains of protein oligomers.

More recent studies reveal that there is greater diversity here than was originally noticed. There are some structures such as in the interchain contacts in the platform of an MHC class II molecule (Reinherz et al., 1999), where each chain contributes one strand at the interface between two different four strand β-sheets to form one 8-strand β-sheet. The association of these chains may be enthalpy driven. A recent examination of 136 homodimeric proteins corroborated the predominance of nonpolar surface but noted that many homo-oligomer subunit interfaces have important hydrophillic aspects (Larsen et

al., 1998). In particular, they found that for homodimers of compact subunits with extended oligomeric interfaces the relative importance of hydrogen bonds, ordered waters and polar surface area to the interfaces was broadly sampled.

Multiple crystal structure analyses: Learning how proteins function

The need to distinguish homo-oligomeric proteins, permanent heterocomplexes, and more transient "non-obligatory" complexes was made explicit in an analysis of protein crystal structures (Jones and Thornton, 1996). Proteins that exist in their biological context physically independent of other proteins may form nonobligate protein-protein complexes by associating with other such proteins (Jones and Thornton, 1996). For example, FcRn in its biological context exists in a form physically independent of other proteins. IgG does also. When these proteins associate their complex is non-obligatory. In contrast, the heavy chain of FcRn does not exist in its biological context in a form independent of the light chain, $\beta 2$ microglobulin . Even though the light chain exists alone in biological systems the complex of the heavy and light chains of FcRn is obligate because a nonobligate complex requires both participants' independent existence. The paramount reason we consider nonobligate complexes distinct is the structures of nonobligate protein-protein complexes are necessarily informative about the function of a protein in a way that the structures of obligate, permanent protein complexes need not be.

Because the crystal structures of nonobligate protein-protein complexes are inherently informative about the function of a protein, analyses of these structures have been helpful in understanding other systematic structure-function studies. The understanding of stoichiometry analyses, binding studies, and the results of site-directed

mutageneses for a given set of molecules are all enhanced when couched in the context of a crystal structure (see for example, Martin and Bjorkman, 1999; Chapter 2). This is true at the scale of multiple sample analyses as well. Analyses of site-directed mutagenesis studies show the fraction of tryptophan to alanine mutations that cause a $\geq$ 2 kcal/mol reduction in binding affinity over the total number of tryptophan mutations is four times greater than the average of the ratio for all mutants (Bogan and Thorn, 1998). Analyses of the structures of nonobligate protein-protein complexes reveal that tryptophan contributes 3.5 times as much to the surfaces of interfaces as it does to the surfaces of proteins as a whole (Lo Conte et al., 1999). It may be that the large size of the effect of the mutation on the binding constant is due to the large steric alteration induced by the mutation. The crystal structure reviews reveal that tryptophans occur in interfaces with a greater frequency than they do on the surface (Lo Conte et al., 1999), which means that size of the effect seen in the mutation studies is not likely the result of undersampling this residue.

The comparison of the review of nonobligate protein-protein complexes and the reviews of mutagenesis studies deepens the understanding of another phenomenon. Mutagenesis studies revealed that alanine substitutions at only 25% of the total number of positions, usually at the center of an interface, generate a sizable reduction in the affinity of the interaction (Cunningham and Wells, 1993; Clackson and Wells, 1995; Bogan and Thorn, 1998). The studies of nonobligate protein-protein complex stuctures reveal several reasons for this. First, because mutageneses fail to affect main-chain atoms they fail to affect one-fifth of the interface surface area that provides two-thirds of the interface hydrogen bonds. Second, three-quarters of the interface area comes from

atoms, almost always those on the interface perimeter, that remain accessible to solvent. The affect of mutations at these residues may be ameliorated by the incorporation of more ordered water molecules (Lo Conte et al., 1999).

Nonobligate complexes allow us to examine the shape of protein surfaces that are functionally significant. One might imagine that rough surfaces are more likely to form specific intimate associations with other rough surfaces. This is the case with the associations between small molecule ligands with proteins. Lewis and Rees developed a metric for the degree of roughness of a protein surface called the fractal dimension (Lewis and Rees, 1985). This is the sensitivity of the size of the surface to the size of the probe used to measure it. Pettit and Bowie, assaying the surface of a protein by its fractal dimension, found that sites involved in binding small molecule ligands were very rough compared to the surface as a whole (Petit and Bowie, 1999). To determine if the interface surface is more or less flat than the surface of a protein generally, it is necessary to extract samples of the rest of the surface. Jones and Thornton first attempted this on a data set including 31 heterocomplexes with a patch-wise analysis of the surface (Jones and Thornton, 1997). Their patch size depended per protein on the size of that protein's interface. They found that in contrast to small molecules, the surfaces of proteins most likely to interact with other proteins are more flat than the surface of proteins are generally.

Another reason to distinguish between obligate and nonobligate protein-protein complexes is that there may be different selection pressures applied to the protein surfaces involved. The protein surfaces involved in nonobligate complexes must be adapted to two environments: aqueous buffer and the interface of the complex. The

surfaces of obligate complexes must be adapted only to the latter. An analysis of a collection of crystal structures has provided a structural insight into this functional question. In a review of 15 protease-protease inhibitor complexes and four antibody-antigen complexes, the interfaces between these proteins contained an average of 10 hydrogen bonds, greater than 1 per 100 Å^2 of buried surface area (Janin and Chothia, 1990). This is more than double the density seen between the subunits of oligomers (Janin et al., 1988), indicating that these interfaces rely to a greater extent on hydrogen bond formation and less on the hydrophobic effect than do the interfaces between subunits of oligomers. A higher density of hydrogen bonds per unit surface area in nonobligate complexes was confirmed in later studies (Jones and Thornton, 1996). This difference is consistent with the proposition that the surfaces involved in the interfaces of nonobligate protein-protein complexes are subjected to selection based on their ability to spend indefinite amounts of time exposed to aqueous buffer and cannot be completely hydrophobic.

In contrast to the differences in the use of hydrogen bonds, the packing of atoms in nonobligate complexes is similar to that found in cores of proteins (Lo Conte et al., 1999). Therefore, while the hydrogen bond formation may be more important to nonobligate complex formation than to the association of oligomer subunits, the importance of van der Waals forces is the similar in both.

The analyses reviewed in the sections above have confirmed the importance of the burial of hydrophobic surface area to the formation of tertiary structure and the association of oligomer subunits (Lee and Richards, 1971; Chothia, 1976). They have also demonstrated the comparatively larger role played by hydrogen bonding in the

formation of nonobligate complexes as compared to permanent obligate complexes (Janin et al., 1988; Janin and Chothia, 1990; Jones and Thornton, 1996) though there are homo oligomers that are exceptions to this rule (Larsen et al., 1988). Multiple crystal structure analyses have indicated that van der Waals interactions are an important (if somewhat under-appreciated) and consistently employed force in protein folding, and oligomer subunit and nonobligate complex formation (Harpaz et al, 1994; LoConte et al., 1999). They have shown that the flat sections of a protein surface are somewhat more likely to bind proteins while rough protein surfaces are more likely to bind small molecule ligands (Jones and Thornton, 1996; Petit and Bowie, 1999). When used in conjunction with analyses of multiple alanine-scanning mutageneses experiments, crystal structure reviews have demonstrated that all residues do not contribute equally to nonobligate protein-protein interactions (Bogan and Thorn, 1998; LoConte et al., 1999). We now review how other analyses aided in finding what characteristics of proteins are common between groups of proteins and what characteristics of proteins may be used to distinguish different groups.

Multiple crystal structure analyses: Distinguishing groups of proteins

Analyses of multiple crystal structures revealed a relationship between the mass of a protein and its surface area that is common to both monomers and oligomers. The algorithm that measures the solvent accessible surface area by Lee and Richards was improved (Shrake and Rupley, 1973) and used to measure the surface area of 46 small monomeric proteins (Janin et al., 1987). This study found that the surface area of these small proteins was related, with some exceptions, to the molecular weight of the proteins by a simple power law. It was proposed that the surface area of oligomeric proteins

would not be related to the mass (Sprang et al., 1979). This proposition was refuted in a study of 23 oligomeric proteins that found the surface area of the protein as well as the area buried between oligomers was a function of the molecular weight of the protein (Miller et al., 1987b). The relationship between the molecular weight and the surface area of the whole oligomeric protein was confirmed in a study of 24 oligomeric proteins (Argos, 1988). However, the relationship between the molecular weight and the amount of surface area buried between oligomers was not found in this study to be a function of the protein's mass. Despite this discrepancy, both studies found that for proteins with the same molecular weight, the surface area of the subunits depended on the number of subunits (Argos, 1988).

Other studies have revealed what else is consistent between groups and what is different. Oligomers and monomers were found to share the same proportions of nonpolar, polar and charged surface area both on their solvent accessible surfaces (57% nonpolar, 23% polar, and 20% charged) and in their cores (58% nonpolar, 33% polar, and 9% charged) (Janin et al., 1988). The surfaces buried between oligomer subunits and those buried in nonobligate protein-protein complexes have amino acid compositions that are more enriched in aromatic amino acids than like the surfaces of proteins generally (Argos, 1988; Janin and Chothia, 1990; Jones and Thornton, 1996; Lo Conte et al., 1999). As noted above, the nonobligate surfaces make more hydrogen bond contacts than occur between the surfaces of subunits in oligomers (Janin and Chothia, 1990; Jones and Thornton, 1996).

As more structures have become available for analyses, it has become possible to distinguish between groups that were previously considered collectively. The initial

studies of homo oligomers treated them as one group of proteins (Janin et al., 1988; Janin and Chothia, 1990; Jones and Thornton, 1996). The more recent work divides homodimers into three classes, the first, largest, class is made of compact subunits with extended interfaces. The relative importance of hydrogen bonds, ordered waters and polar surface area to the homo oligomer interfaces is broadly sampled within this class. The other two classes, compact subunits with compact interfaces and a class made up of intertwined interfaces rely on hydrophobic cores and extensive hydrophobic surface area, respectively (Larsen et al., 1998). Early studies of nonobligate protein-protein complexes found the complexed proteins underwent little conformational change when binding each other (Janin and Chothia, 1990; Jones and Thornton, 1996). These older, smaller studies saw little movement of the proteins upon complex formation and that result was confirmed in a larger study, for small proteins. A recent review of 75 nonobligate protein-protein complexes (Lo Conte et al., 1999) noticed that for larger proteins with interfaces burying between 2000-4700 Å^2 total this is not always the case. Among others, the binding of $G_{i\beta\gamma}$ with phosducin, the binding of $G_{i\alpha}$ to transducin and CDK2 binding to cyclin A are all accompanied by large changes in the structure. These structural changes are of three general types, disorder to order transitions, large main-chain movements (primarily alterations in loop structure), and changes in the positions of the domains relative to each other (Lo Conte et al., 1999).

It has been shown that the core of a monomer is more like the core of an oligomer than it is like the surface of a monomer (Janin et al., 1988). Oligomers and monomers also share a general relationship between the mass of a protein and its surface area (Miller et al., 1987b; Janin et al., 1988; Argos, 1988). While nonobligate complexes have

more hydrogen bonds between the interface surfaces than occur in obligate complexes, they share a similar propensity for aromatic amino acids as well as efficient packing (Janin et al., 1988; Argos, 1988). The distinction between permanent, obligate complexes and transient, nonobligate ones as well the difference between small, more elastic, nonobligate complexes, and larger inelastic ones was formalized in these analyses (Jones and Thornton, 1996; LoConte et al., 1999). To conclude the review of multiple crystal structure analyses, we will examine the tools that have made these analyses possible.

Multiple crystal structure analyses: Tools for quantifying aspects of structures

We have observed above the importance of the algorithms for calculating the basic physical properties of surface area and volume to analyzing collections of crystal structures. We have seen how the output of these algorithms may be partitioned in a variety of ways to tell us much about the character of the examined proteins. There are aspects of nonobligate protein-protein complexes in particular that are captured best by algorithms that measure more subtle aspects of a protein's shape.

The shape of the interface of a nonobligate protein-protein complex has been described quantitatively by two numbers in a analyses of multiple crystal structures. In a review of 32 homo-oligomers, 4 obligate hetero-oligomers, 10 enzyme-inhibitor complexes, 4 antibody-antigen complexes and 7 other nonobligate complexes, the planarity and the circularity of both interface surfaces were measured (Jones and Thornton, 1996). The atoms that bury surface area to form the interface were used to define the absolute shape of an interface. These atoms are dispersed in three dimensions such that they do not form any regular, easily defined shape. To describe this shape, the

best-fit plane through the dispersion of atoms is found. Jones and Thornton employ this method and describe the way the dispersion fits a plane with two values, the planarity and the circularity. The planarity is the root mean squared deviation away from the plane, the smaller this value the more planar the dispersion. The circularity is the ratio of what may be thought of as the two axes of the plane. If this ratio is equal to 1.0, the dispersion of atoms that make up the interface will appear circular if one views the interface from a perspective normal to the best-fit plane. The smaller this ratio, the more oblong the interface. Using this tool to compare the interfaces of obligate and nonobligate complexes revealed that nonobligate interfaces were more planar than but similarly circular to obligate complexes.

Jones and Thornton also measured the degree to which the shape of one interface surface matched the other, or how well the proteins "fit together." They did this by measuring the void volume between the complexed proteins. Larger volumes corresponded to poorer fits. Obligate complexes and enzyme-inhibitor complexes fit together well, while antibody-antigen complexes fit together less well (Jones and Thornton, 1996). This confirmed the results arrived at by a different algorithm that measures the degree to which proteins fit in a different way (Lawrence and Colman, 1993). To assess the shape complementarity of two proteins in complex, this algorithm measures both the distance between the surfaces of atoms as well as the angle between the normals extending from the surface. Perfectly complementary surfaces have a shape complementarity of 1.0, while completely unlike surfaces register 0.0 shape complementarity. When this algorithm was applied to the structures of four protease-protease inhibitor complexes, five obligate complexes, and six antibody antigen

complexes, the first two groups were found to have a higher degree of shape complementarity than the latter.

The shape of each of the surfaces in the interface of a nonobligate protein-protein complex has been measured as well as the degree of fit between them. These values have been found useful in describing differences between the biologically distinguishable types of nonobligate protein-protein complexes (Jones and Thornton, 1996; Lawrence and Colman, 1993).

Conclusions

Analyses of collections protein crystal structures have provided original insights into protein folding. Analyses of the types of protein surface buried illuminated the role of hydrogen bond dominated secondary structure formation and hydrophobic surface burial dominated tertiary structure formation (Lee and Richard, 1971; Chothia, 1976). Analysis of the volumes of atoms in cores of proteins in high-resolution crystal structures elucidated the importance of van der Waals forces on the packing of atoms in a folded protein (Harpaz et al., 1994). This result implies that the solution-transfer model, which correctly captures the hydrophobic effect, does not fully account for the role of atom packing in protein folding (Kauzmann, 1987; Harpaz et al., 1994). The indication of these studies is that dissolution of cyclic dipeptides crystals in water may be a better model for protein folding than the broadly applied solution transfer model.

Analyses of multiple crystal structures help us understand how separate polypeptide chains associate. The low number of hydrogen bonds between the subunits of oligomers and the high density of the participating amino acids suggest that the roles of hydrogen bonding, van der Waals forces, and the hydrophobic effect in the formation

of quaternary structure are similar to those found in forming tertiary structure (Janin et al., 1988; Harpaz et al., 1994). The greater number of hydrogen bonds and similar amino acid densities at the interfaces of nonobligate protein-protein complexes indicates that these complexes require the tight packing of protein cores but rely less on the hydrophobic effect and more on enthalpic contributions than do more permanent associations (Janin and Chothia, 1990; LoConte et al., 1999). These studies also reveal a prominent role played by aromatic amino acids at interfaces (Argos, 1988; LoConte et al., 1999).

The analyses of multiple crystal structures helped define how we group protein crystal structures and demonstrated how the different groups are distinct and how they are similar. Analyses of monomeric and oligomeric proteins revealed that in most respects measurable in collective studies these proteins are similar (Miller et al., 1987b; Janin et al., 1988). In contrast, studies of nonobligate protein-protein complexes have found that within this category there are differences between the extent conformational change induced by binding and in the degree of fit between the complexed proteins (Lawrence and Colman, 1993; Jones and Thornton, 1996; LoConte et al., 1999). These distinctions indicate that this class may need to be subdivided further.

References

Argos, P. (1988) An investigation of protein subunit and domain interfaces. *Prot. Eng.* 2, 101-113.

Baldwin, R. L. (1986) Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl.Acad. Sci. USA*, 83, 8069-8072.

Bogan, A. A. and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces. *J. Mol. Biol.*, 280, 1-9.

Brandts, J. F., Oliveira, R., J., and Westort, C. (1970) Thermodynamics of protein denaturation. Effect of pressure on denaturation of ribonuclease A. *Biochemistry* 9, 1038-1047.

Chen, Z. and Bode, W. (1983) Refined 2.5 Å X-ray crystal structure of the complex formed by porcine kallikrein A and the bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 164, 283-311.

Chothia, C. (1975) Structural invariants in protein folding. *Nature*, 254, 304-308.

Chothia, C. (1976) The *Nature* of accessible and buried surfaces in proteins. *J. Mol. Biol.*, 105, 1-12.

Chothia, C. and Janin, J. (1975) Principles of protein-protein recognition. *Nature*, 256, 705-708.

Clackson, T. and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science*, 267, 383-386.

Cunningham, B. C. and Wells, J. A. (1993) Conparison of a structural and functional epitope. *J. Mol. Biol.* 234, 554-563.

Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E. and Wiley, D. C. (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature*, 384, 134-141.

Gerstein, M., Tsai, J. and Levitt, M. (1995) The volume of atoms on the protein surface: Calculated from simulation, using Voronoi polyhedra. *J. Mol. Biol.*, 249, 955-966.

Harpaz, Y., Gerstein, M. and Chothia, C. (1994) Volume changes on protein folding. *Structure*, 2, 641-649.

Hawley, S. A. (1971) Reversible pressure-temperature denaturation of chymotrypsinogen. *Biochemistry* 10, 2436-2442.

Janin, J. and Chothia, C. (1990) The structure of protein-protein recognition sites. *J. Biol. Chem.*, 265, 16027-16030.

Janin, J., Miller, S. and Chothia, C. (1988) Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.*, 204, 155-164.

Murphy, K. P. and Gill, S. J. (1991) Solid model compounds and the thermodynamics of protein unfolding. *J. Mol. Biol.* 222, 699-709.

Jones, S. and Thornton, J. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U S A*, 93, 13-20.

Kauzmann W. (1959) Some factors in the interpretation of protein denaturation. Advan. Protein Chem. 14, 1-63.

Kauzmann W. (1987) Thermodynamics of unfolding. *Nature*, 325, 763-764.

Lawrence, M. C. and Colman, P. (1993) Shape complementarity at protein/protein interfaces. *J. Mol. Biol.*, 234, 946-950.

Larsen, T. A., Olson, A. J., and Goodsell, D. S. (1998) Morphology of protein-protein interfaces. *Structure*, 6, 421-427.

Lee, B. K. and Richards, F. M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55, 379-400.

Lewis, M. and Rees, D. C. (1985) Fractal surfaces of proteins. *Science*, 230, 1163-1164.

LoConte, L., Chothia, C. and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.*, 285, 2177-2198.

Miller, S., Lesk, A. M., Janin, J. and Chothia, C. (1987) The accessible surface area and stability of oligomeric proteins. *Nature*, 328, 834-835.

Miller, S., Janin, J., Lesk, A. M. and Chothia, C. (1987) Interior and surface of monomeric proteins. *J. Mol. Biol.* 196, 641-656.

Pauling, L., Corey, R. B. and Bramson, H. R. (1951) The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U S A*, 37, 205-211.

Pauling, L. and Corey, R. B. (1951) Configurations of polypeptide chains with favored orientations around single bonds: two pleated sheets. *Proc. Natl. Acad. Sci. USA*, 37, 729-740.

Pauling, L. and Corey, R. B. (1951) The pleated sheet, a new layer configuration of polypeptide chains. *Proc. Natl. Acad. Sci. USA*, 37, 251-261.

Pettit, F. K. and Bowie, J. U. (1999) Protein surface roughness and small molecular binding sites. *J. Mol. Biol.*, 285, 1377-1382.

Richards, F. M. (1974) The interpretation of protein structures: total volume, group volume distributions and packing density. *J. Mol. Biol.* 82, 1-14.

Reinherz, E. L., Tan, K., Tang, L., Kern, P., Liu, J., Xiong, Y., Hussey, R. E., Smolyar, A., Hare, B., Zhang, R., Joachimiak, A., Chang, H. C., Wagner, G. and Wang, J. (1999) The crystal structure of a T-Cell receptor in complex with peptide and MHC Class II. *Science*, 286, 1913-1921.

Schindelin, H., Kisker, C., Schlessman, J. L., Howard, J. B. and. Rees, D. C. (1997) Structure of ADP x AIF4(-)-stabilized nitrogenase complex and its implications for signal transduction. *Nature* 387, 370-376.

Shrake, A. and Rupley, J. A. (1973) Environment and exposure to solvent protein atoms Lysozyme and Insulin. *J. Mol. Biol.* 79, 351-371.

Sprang, S. Yang, D. and. Fletterick, R. J. (1979) Solvent accessibility properties of complex proteins. *Nature* 280, 333-335.

Stites, W. E. (1997) Protein-protein interactions: interface structure, binding thermodynamics, and mutational analysis. *Chem. Rev.*, 97, 1233-1250.

Zipp, A. and Kauzmann, W. (1973) Pressure denaturation of metmyoglobin. *Biochemistry*, 12, 4217-4228.

Chapter 2:

Characterization of the 2:1 Complex between the Class I MHC-Related Fc Receptor and Its Fc Ligand in Solution

This chapter describes the generation of the heterodimeric Fc molecule and the characterization of its interaction with FcRn.

Characterization of the 2:1 Complex between the Class I MHC-Related Fc Receptor and Its Fc Ligand in Solution[†]

W. Lance Martin and Pamela J. Bjorkman*

Division of Biology 156-29 and Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125

ABSTRACT: The neonatal Fc receptor (FcRn) facilitates the transfer of maternal immunoglobulin G (IgG) to offspring and prolongs the half-life of serum IgG. FcRn binds IgG in acidic intracellular vesicles and releases IgG upon exposure to the basic pH of the bloodstream. The crystal structure of an FcRn/Fc complex revealed FcRn dimers bridged by homodimeric Fc molecules to create an oligomeric array with two receptors per Fc [Burmeister et al. (1994) *Nature 372*, 379–383], consistent with the 2:1 FcRn:Fc stoichiometry observed in solution [Huber et al. (1993) *J. Mol. Biol. 230*, 1077–1083; Sánchez et al. (1999) *Biochemistry 38*, 9471–9476]. Two distinct 2:1 FcRn/Fc complexes were present in the cocrystal structure: a complex containing an FcRn dimer interacting with an Fc and a complex in which single FcRn molecules are bound to both sides of the Fc homodimer. To determine which of the two possible 2:1 FcRn/Fc complexes exists in solution, we generated recombinant Fc molecules with zero, one, and two FcRn binding sites and studied their interactions with a soluble form of rat FcRn. The wild-type Fc with two FcRn binding sites binds two FcRn molecules under all assay conditions, and the nonbinding Fc with no FcRn binding sites shows no specific binding. The heterodimeric Fc with one FcRn binding site binds one FcRn molecule, suggesting that the 2:1 FcRn/wild-type Fc complex formed in solution consists of single FcRn molecules binding to both sides of Fc rather than an FcRn dimer binding to a single site on Fc.

The neonatal Fc receptor (FcRn)[1] transports immunoglobulin G (IgG) across epithelial cell barriers. FcRn was originally discovered in the intestine of newborn rodents (reviewed in ref *1*), where it transfers maternal immunoglobulin in ingested milk to the bloodstream of the newborn, thereby allowing passive immunization of the neonate to antigens encountered by the mother. More recently, FcRn has been characterized in adult animals. FcRn in human placenta is thought to transport maternal IgG to the fetus, and recent evidence suggests that FcRn functions throughout life to rescue serum IgG from degradation (reviewed in refs *1–3*). In all of its functions, FcRn binds IgG at acidic pH (≤6.5) in intracellular transport vesicles and releases IgG at the basic pH of the blood (pH ~7.4).

FcRn is a type I membrane glycoprotein consisting of an extracellular domain that resembles class I MHC molecules and a short (43 residue) cytoplasmic tail (*4*). Most biochemical and structural analyses of the FcRn/Fc interaction have been done using a soluble form of the extracellular portion of rat or mouse FcRn, which is composed of the heavy chain extracellular domains (α1, α2, and α3) bound to β2-microglobulin. The crystal structures of soluble FcRn alone (*5*) and of an FcRn/Fc complex (*6*) revealed a dimeric arrangement of receptors. Binding studies involving FcRn mutants with alterations at the dimer interface demonstrated that receptor dimerization is required for high-affinity binding of IgG (*7, 8*). We therefore suggested that the crystallographically observed dimers represent dimers induced by ligand binding when FcRn is tethered to a membrane (*9*). In the cocrystals, the FcRn dimers are bridged by Fc molecules such that each of the two potential FcRn binding sites on Fc interacts with one of the FcRn molecules in the receptor dimer, resulting in a long "oligomeric ribbon" in which there are two receptors for every Fc dimer (*6*). At micromolar concentrations in solution, there is no detectable formation of the oligomeric ribbon (*10*). Instead, purified FcRn/Fc complexes consist of three molecules: two receptors and one Fc, which presumably represent a portion of the 2*n*:*n* ribbon found in the crystals (*10, 11*). There are two distinct 2:1 complexes in the cocrystal structure that could account for the FcRn/Fc complex that forms in solution (Figure 1). In one, Fc binds to an FcRn dimer using one of its two potential FcRn binding sites (left shaded portion), and in the other, single FcRn molecules bind to both sides of Fc (right shaded portion). Computational studies suggested that Fc bound to FcRn is bent so as to more optimally contact the FcRn dimer (*12*). A bent, rather than symmetrical, structure of Fc bound to FcRn is compatible with the low-resolution cocrystal structure since the hinge-proximal portions of the C_H2 domain were disordered (*6*). If Fc is distorted when bound

¹ Abbreviations: CHO, Chinese hamster ovary; Fc, Fc fragment from immunoglobulin G; FcRn, Fc receptor, neonatal; hdFc, heterodimer Fc; IgG, immunoglobulin G; K_D, equilibrium dissociation constant; nbFc, nonbinding Fc; RU, resonance units; wtFc, wild-type Fc.

2:1 FcRn/Fc complex
with FcRn dimer

2:1 FcRn/Fc complex
with bridging Fc
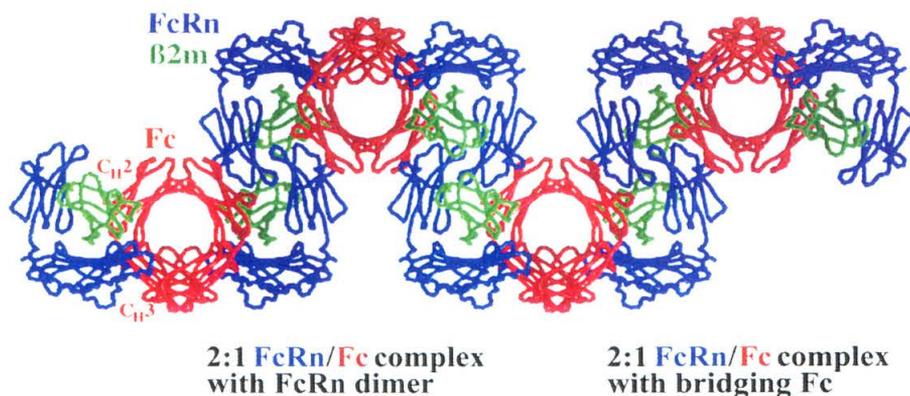
FIGURE 1: FcRn/Fc complexes in the 2n:n oligomeric ribbon observed in the FcRn/Fc cocrystals. FcRn dimers are bridged by homodimeric Fc's. The shaded portions represent the two different 2:1 FcRn/Fc complexes that can be extracted from the ribbon (see text). Although the Fc is depicted as being 2-fold symmetric, the low-resolution FcRn/Fc cocrystal structure did not give information about the location of the hinge-proximal portions of the C_H2 domains since these were disordered (6). Computational studies suggest that Fc bends in response to contacting an FcRn dimer (12); thus either of the depicted 2:1 FcRn/Fc complexes is a possibility for the 2:1 complex that forms in solution. If the oligomeric ribbon network forms under physiological conditions, each FcRn dimer would be associated with a membrane parallel to the plane of the paper: the left-most dimer is associated with a membrane below the plane of the paper, the central dimer is associated with a membrane above the paper, and the right-most dimer is again associated with the membrane below the plane of the paper.

to an FcRn dimer, a second FcRn might be prevented from binding to the other Fc polypeptide chain in solution. Thus either of the 2:1 FcRn/Fc complexes shown in Figure 1 are a possibility for the complex that forms in solution. To distinguish which 2:1 complex forms in solution, we constructed Fc proteins, similar to those previously expressed by Ward and colleagues (2, 13−18), that contain zero, one, or two functional FcRn binding sites and studied their binding to soluble FcRn.

MATERIALS AND METHODS

Construction of Fc Expression Vectors. A rat IgG2a cDNA (kind gift of Mark Agila, University of California, Davis) was modified to encode a secreted Fc fragment as follows: the DNA encoding the V_H and C_H1 domains was removed using loop-out mutagenesis (19) to generate an in-frame fusion between the secretion signal sequence and the hinge region. The complete construct encodes the signal sequence fused to IgG2a residues 223−447 (EU numbering; 20), which corresponds to the hinge, C_H2, and C_H3 domains of wild-type Fc (wtFc). The nonbinding Fc (nbFc) construct was generated from the wtFc construct by introducing mutations at the codons for some of the amino acids previously shown to be critical for FcRn binding (2, 9, 13−18, 21) to make the following substitutions: Thr252 to Gly, Ile253 to Gly, Thr254 to Gly, His310 to Glu, His433 to Glu, and His435 to Glu. PCR was used to add a factor Xa cleavage site and 6×-His tag to the 3′ end of the nbFc construct (added sequence encodes the following residues C-terminal to residue 447: Gly-Ile-Glu-Gly-Arg-Gly-Ser-Ser-His-His-His-His-His-His). The wtFc and nbFc constructs were subcloned after sequencing into the mammalian cell expression pBJ5-GS (22), which carries the glutamine synthetase gene as a means of selection and amplification in the presence of the drug methionine sulfoximine (23).

Expression of Fc Proteins. Chinese hamster ovary (CHO) cells were cotransfected with the wtFc and nbFc expression vectors, and selection and amplification of stable cell lines

using methionine sulfoximine were carried out as described (22, 24). Successfully transfected cells should secrete a mixture of wtFc and nbFc homodimers and heterodimeric Fc (hdFc) composed of one wtFc and one nbFc polypeptide chain. Cell lines secreting wtFc and hdFc were identified by precipitation at pH 6.0 of [^{35}S]methionine/cysteine (ICN Pharmaceuticals, Inc.) metabolically labeled supernatants using soluble FcRn coupled to CNBr-activated Sepharose beads (Gibco-BRL) (11). Bound Fc was eluted from the FcRn-coupled beads by raising the pH to 8.0 and loaded onto a 10% SDS−PAGE gel run under reducing conditions. Bands migrating with apparent molecular masses of 30 and 31 kDa were visualized using a PhosphorImager screen (Molecular Dynamics) (data not shown). The lower band was identified as wtFc by comparison with the Fc protein expressed in cells transfected with the wtFc vector alone (see below). The upper band corresponds to nbFc, which migrates more slowly than wtFc due to addition of the factor Xa site and 6×-His tag to its C-terminus. The majority of the labeled protein migrated as the 30 kDa band, indicating that the FcRn-coupled beads precipitated wtFc homodimers and hdFc. After addition of glycerol to 10%, NaCl to 300 mM, and imidazole to 10 mM, labeled supernatants were also precipitated with Ni−NTA superflow agarose beads (Qiagen). Bound proteins were eluted from the Ni−NTA beads by addition of 1 M imidazole and reducing sample buffer and visualized after SDS−PAGE as described above. The 30 and 31 kDa bands were again present, with the 31 kDa band in excess in this instance, indicating that the nickel beads precipitated His-tagged nbFc homodimers and hdFc.

CHO cells were also transfected with the wtFc expression vector alone. After selection and amplification, cells expressing wtFc homodimers were identified using FcRn-coupled beads as described above.

Purification of Fc Proteins. Secreted wtFc homodimers were isolated from supernatants of CHO cells transfected with only the wtFc expression vector using a modification of a previously described functional purification involving

pH-dependent binding to FcRn immobilized on Sepharose beads (*11*). The pH of the harvested growth media was adjusted to 5.8 with 1 M sodium cacodylate, pH 5.5, (~50 mL/500 mL harvest) then passed over a 10 mL FcRn−Sepharose column at 0.5 mL/min. After washing with 200 mL of 50 mM sodium cacodylate, pH 5.5, 150 mM NaCl, wtFc was eluted with 50 mM Tris-Cl, pH 8.0, 150 mM NaCl, then concentrated, and exchanged into 50 mM Tris-Cl, 50 mM bis-Tris-propane-Cl, pH 8.0. wtFc was loaded onto a Uno-Q1 anion-exchange column (Bio-Rad) mounted on a Biocad 700E perfusion chromatography system (Perkin-Elmer) at 5 mL/min. The Uno-Q1 column was equilibrated with the FcRn loading buffer and then subjected to a pH gradient from 7.5 to 6.0. Under these conditions, wtFc (calculated pI = 7.1) (*25*) does not bind to the column, allowing it to be separated from contaminants that bind. wtFc was concentrated, then purified, and exchanged into 20 mM sodium phosphate, pH 6.0, 150 mM NaCl by flowing it over a Superdex 200 HR 10/30 gel filtration column (Pharmacia) at 0.3 mL/min.

hdFc and nbFc were purified from CHO cells secreting a mixture of wtFc, hdFc, and nbFc. Since only the nbFc polypeptide chain carries a 6×-His tag, hdFc and nbFc bind to a nickel column, whereas wtFc flows through. hdFc and nbFc can then be separated from each other using the FcRn affinity column, which binds hdFc but not nbFc. Harvest media were dialyzed twice using 6000−8000 Da Spectra/Por membranes (Spectrum) against 10 volumes of 20 mM Tris-Cl, pH 8.0, 150 mM NaCl, and 0.05% NaN$_3$ to remove a media component that stripped nickel from the Ni−NTA column. The harvest media were then supplemented with glycerol to 10%, NaCl to 300 mM, and imidazole to 10 mM. The media were passed over a 20 mL Ni−NTA superflow agarose column at 0.5 mL/min and washed with 200 mL of 50 mM Tris-Cl, pH 8.0, 10% glycerol, 300 mM NaCl, 10 mM imidazole, 0.05% NaN$_3$. Nickel binding proteins (hdFc and nbFc) were eluted from the column with 50 mM Tris-Cl, pH 8.0, 10% glycerol, 300 mM NaCl, 250 mM imidazole, 0.05% NaN$_3$. The eluent was concentrated and exchanged into 20 mM sodium cacodylate, pH 5.5, 150 mM NaCl and passed over an FcRn−Sepharose column, which was eluted by raising the pH to 8.0. To avoid potential contamination of the hdFc with wtFc, different FcRn−Sepharose columns were used for purifying wtFc and hdFc. After passing the flow-through over the FcRn−Sepharose column again, the flow-through (nbFc) and the eluent (hdFc) were each loaded onto the Uno-Q1 anion-exchange column as described for wtFc. The hdFc (calculated pI = 6.9) (*25*) binds to the column and elutes at pH 7.4, 6 mL after the wtFc peak. The nbFc (calculated pI = 6.7) (*25*) binds to this column and elutes at pH 7.1, 10 mL after the hdFc peak. In addition to enhancing purity of the samples, this chromatography facilitates evaluation of the degree to which each Fc is free of other contaminating Fc species. Following anion exchange, nbFc and hdFc were concentrated and further purified by gel filtration on a Superdex 200 HR 10/30 column as described for wtFc.

Final yields for the purified Fc proteins ranged from 3 to 7 mg/L of harvest.

Expression and Purification of Soluble FcRn. The FcRn used for these studies was the previously described soluble form of the rat FcRn extracellular region (residues 1−269

of the mature FcRn heavy chain bound to rat β2-micro-globulin) produced in CHO cells (*22*). FcRn was isolated from harvested growth media using a functional purification involving pH-dependent binding to an IgG affinity column (*22*). The protein was further purified on Uno-Q1 anion-exchange and Superdex 200 HR 10/30 gel filtration columns and exchanged into assay buffer as described for wtFc.

Determination of Protein Concentrations. FcRn and Fc concentrations were determined spectrophotometrically using extinction coefficients at 280 nm of 84 900 M^{-1} cm^{-1} (FcRn) and 60 900 M^{-1} cm^{-1} (wtFc, hdFc, and nbFc). Extinction coefficients that are valid for denatured protein were first calculated from the protein sequences as described (*26*); then A_{280} measurements for a fixed amount of each protein were compared in 6.0 M GuHCl and aqueous solutions, and the coefficient was adjusted if necessary.

Coprecipitation of FcRn and Fc. Fc proteins were analyzed for their ability to bind simultaneously to more than one FcRn using a modification of a previously described column binding assay (*11*). For each reaction, 20 μL of FcRn−Sepharose beads was washed and resuspended in 50 μL of sodium phosphate, pH 6.0, 150 mM NaCl in a 1.7 mL Eppendorf tube and then incubated with 20 μg of wtFc, nbFc, or hdFc in ~5 μL or with 5 μL of buffer. After washing twice, the beads were resuspended in 50 μL of the same buffer including 20 μg of soluble FcRn. After two subsequent washes, the proteins bound to the beads were eluted with 16 μL of 1 M Tris-Cl, pH 8.0. Eluted proteins were run on a 10% SDS−PAGE gel under reducing conditions and stained with Coomassie brilliant blue.

Nonequilibrium Gel Filtration. The stoichiometry of FcRn/Fc complexes was determined using conventional gel filtration chromatography under nonequilibrium conditions as described (*10*). FcRn was incubated with the various Fc's at molar ratios between 3:1 and 1:1 in 20 mM sodium phosphate, pH 6.0, 150 mM NaCl, keeping the concentration of Fc fixed at 10 μM. After 20 min at room temperature, 25 μL was injected onto a Superdex 200 PC 3.2/30 gel filtration column (Pharmacia) equilibrated in the sodium phosphate buffer, which was run at 0.1 mL/min using a SMART micropurification system (Pharmacia). The absorbance of the eluent was monitored at 280 nm, and fractions were analyzed by SDS−PAGE (data not shown).

Equilibrium Gel Filtration. The equilibrium column chromatography method of Hummel and Dreyer (*27*) and a SMART micropurification system were used to analyze the association of FcRn with the Fc proteins at equilibrium as previously described (*10*). A Superdex 200 PC 3.2/30 gel filtration column was equilibrated with and run in 20 mM sodium phosphate, pH 6.0, 150 mM NaCl containing 2 μM FcRn (equilibration buffer) at 0.1 mL/min. Samples (20 μL) including a 2 μM amount of one of the Fc's and various concentrations of FcRn were incubated for 20 min at room temperature in equilibration buffer that contained 2 μM FcRn. Samples were injected onto the column, and the absorbance of the eluent was monitored at 280 nm.

Biosensor Assays. A BIAcore 2000 biosensor system (Pharmacia, LKB Biotechnology) was used to assay the interaction of FcRn with the Fc molecules. This system includes a biosensor chip with a dextran-coated gold surface to which one protein (referred to as the "ligand") is covalently immobilized. Binding of an injected protein (the "analyte")

to the immobilized protein results in changes that are directly proportional to the amount of bound protein and read out in real time as resonance units (RU) (28, 29). FcRn or one of the Fc's was covalently immobilized to three of the four flow cells on a CM5 biosensor chip (Pharmacia) using standard primary amine coupling chemistry (BIAcore manual). Each protein was coupled at three different densities (~200, ~400, and ~1500 RU), and the fourth flow cell was mock coupled using buffer to serve as a blank. For deriving kinetic constants, we used binding experiments conducted for short times (30 s) using fast flow rates (100 μL/min) over flow cells coupled at low density (~200 or ~400 RU). These conditions were chosen to minimize mass transport effects upon the kinetics of binding reactions (30). Kinetic constants were derived from the sensorgram data using BIAevaluation version 3.0, which simultaneously fits the association and dissociation phases of the sensorgrams and globally fits all curves in the working set. Sensorgrams were fit to models supplied by the BIAevaluation 3.0 package: the "Langmuir binding" model (a single class of noninteracting binding sites in a 1:1 binding interaction), the "heterogenenous ligand" model (two or more populations of noninteracting binding sites), and the "bivalent analyte" model (the injected protein can bind to two immobilized proteins) (see Figure 5 legend). The appropriate model was chosen on the basis of the quality of the fit to the data, the robustness of the fit under different experimental conditions, and consistency between the binding model and structural information regarding the binding mechanism. Equilibrium dissociation constants (K_D's) were derived from the ratios of rate constants ($K_D = k_d/k_a$, where k_a and k_d are the association and dissociation rate constants, respectively). For some of the binding interactions, we also derived K_D's using an equilibrium-based approach that is not affected by mass transport effects. In these experiments, binding reactions were allowed to closely approach or to reach equilibrium. K_D's were derived by nonlinear regression analysis of plots of R_{eq} (the equilibrium binding response) versus the log of the analyte concentration. The fit of data to binding models assuming a bivalent analyte (A. P. West, unpublished results) or to one or more classes of noninteracting binding sites on the ligand was examined, and the appropriate model was chosen as described (31). For each analysis, the bulk refractive index parameter was set to zero for every concentration.

RESULTS

wtFc and hdFc, but Not nbFc, Bind to FcRn. In previous studies by Ward and colleagues, recombinant versions of wild-type (two FcRn binding sites), nonbinding (zero FcRn binding sites), and heterodimeric (one FcRn binding site) mouse Fc were produced in bacteria (13−15). These proteins were used for in vivo catabolic and transcytosis studies, as well as for biochemical analyses of binding to soluble mouse FcRn (13−15, 17, 18, 32, 33). The heterodimeric Fc was shown to bind mouse FcRn (33), but it was not protected from serum degradation or transported across the mouse intestine as efficiently as wild-type Fc (13, 14). We expressed analogous versions of rat Fc in stably transfected CHO cells in order to generate milligram quantities of glycosylated Fc fragments that could be used for crystallographic and biochemical studies involving soluble FcRn.
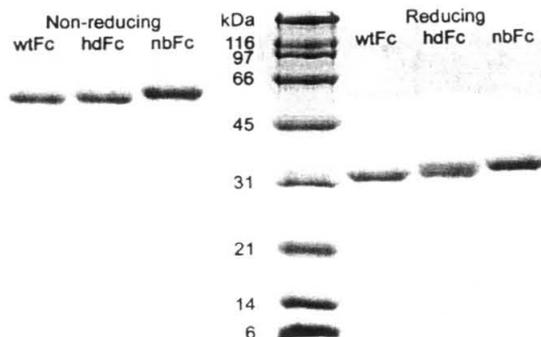
FIGURE 2: SDS−PAGE analysis of Fc proteins. Samples were run under reducing or nonreducing conditions on a 10% SDS−PAGE gel.

CHO cells were cotransfected with expression vectors encoding a secretion signal sequence followed by the hinge, C_H2, and C_H3 domains of wtFc and nbFc derived from a rat IgG2a gene. The nbFc construct was generated from the wtFc construct by introduction of a C-terminal 6×-His tag sequence and incorporation of substitutions identified previously that reduce or eliminate binding of Fc to FcRn (2, 9, 13−18, 21). FcRn- and nickel-based precipitation methods were used to identify transfected cells that secrete a mixture of wtFc homodimers, nbFc homodimers, and hdFc molecules (data not shown).

The purification procedures used to isolate the Fc's require that wtFc and hdFc bind to FcRn at pH 6 but not pH 8, as observed for IgG (34), and that nbFc be unable to bind to FcRn. hdFc and nbFc were purified from the harvested growth media of stably cotransfected cells using a combination of nickel and FcRn−Sepharose chromatography. Supernatants were first passed over a Ni−NTA column, which binds nbFc and hdFc. The eluted proteins were then run over an FcRn−Sepharose column at pH 6 to separate nbFc and hdFc. nbFc was further purified from the flow-through of this column, whereas hdFc was purified after elution at pH 8 from the FcRn−Sepharose column. To obtain large quantities of wtFc, this protein was purified from the harvested growth media of cells transfected with the wtFc expression vector only, using pH-dependent binding to the FcRn−Sepharose column. SDS−PAGE analysis of the purified Fc's under reducing conditions revealed single bands at the expected molecular weights for nbFc and wtFc and two bands corresponding to the nbFc and wtFc polypeptide chains for the hdFc (Figure 2). Under nonreducing conditions, each protein migrates as a dimer, demonstrating that the hinge region interchain disulfide bonds had formed correctly (Figure 2). N-Terminal sequence analysis of purified hdFc revealed a single amino acid sequence (Val-Pro-Arg-Glu-X-Asn-Pro-X-Gly-X, where X corresponds to cysteine, which was not determined using this protocol) (data not shown). This sequence corresponds to residues 223−232 of Fc, demonstrating that the secretion signal sequence had been properly cleaved from the wtFc and nbFc polypeptide chains.

wtFc and hdFc Show Different Properties When Binding to Immobilized FcRn. We previously used a column binding assay to show that more than one FcRn molecule can bind to purified rat Fc (11). In this experiment, soluble rat FcRn
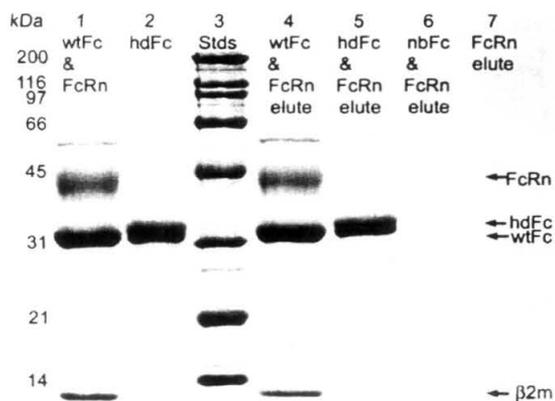
FIGURE 3: SDS−PAGE analysis of proteins eluted from immobilized FcRn. Samples were run under reducing conditions on a 10% SDS−PAGE gel. Lanes 1 and 2: 5 μg of each of the indicated proteins. Lanes 4−7: Proteins eluted from FcRn−Sepharose when incubated with the following proteins (20 μg each): wtFc and FcRn (lane 4), hdFc and FcRn (lane 5), nbFc and FcRn (lane 6), and FcRn (lane 7). Only wtFc is able to bind simultaneously to immobilized FcRn and added soluble FcRn (lane 4). hdFc bound to immobilized FcRn does not bind additional FcRn molecules (lane 5). nbFc and FcRn (lane 6) or FcRn alone (lane 7) does not bind to immobilized FcRn.

was covalently coupled to Sepharose beads and used to bind Fc at pH 6.0. When additional FcRn was passed over the column, it bound to the FcRn/Fc complex, indicating that more than one FcRn molecule could bind to a single Fc (*11*). In the present study, we performed a precipitation-based version of this experiment using wtFc and hdFc. As demonstrated previously (*11*), wtFc is capable of binding simultaneously to immobilized FcRn on the beads and to soluble FcRn, but no detectable soluble FcRn binds to immobilized FcRn in the absence of added Fc (Figure 3). By contrast, hdFc binds to immobilized FcRn but does not bind additional soluble FcRn, suggesting that hdFc can only bind to a single FcRn under the conditions of this experiment. As expected, no detectable nbFc bound to the immobilized FcRn (Figure 3).

Gel Filtration Analyses Demonstrate Different Stoichiometries for wtFc and hdFc Binding to FcRn. We previously described nonequilibrium and equilibrium gel filtration assays to determine the stoichiometry of rat and mouse FcRn complexes with Fc (*10*). In the present study, we compared the properties of wtFc, hdFc, and nbFc in these assays. In the nonequilibrium-based experiments, various ratios of FcRn and one of the Fc's were incubated at pH 6.0, and the FcRn/Fc complex was separated from the free proteins on a gel filtration column. There was no detectable complex formed when mixtures of nbFc and FcRn were chromatographed together on the column (data not shown). For wtFc, we obtained results similar to those previously reported (*10*), such that virtually all of the protein chromatographs as the complex at a 2:1 molar ratio of FcRn to Fc (Figure 4A). When the input ratio of FcRn to wtFc is greater than 2:1, there is an additional peak corresponding to free FcRn; when the input ratio of FcRn to wtFc is less than 2:1, there is an additional peak corresponding to free wtFc (verified by SDS−PAGE; data not shown). By contrast, for hdFc, the
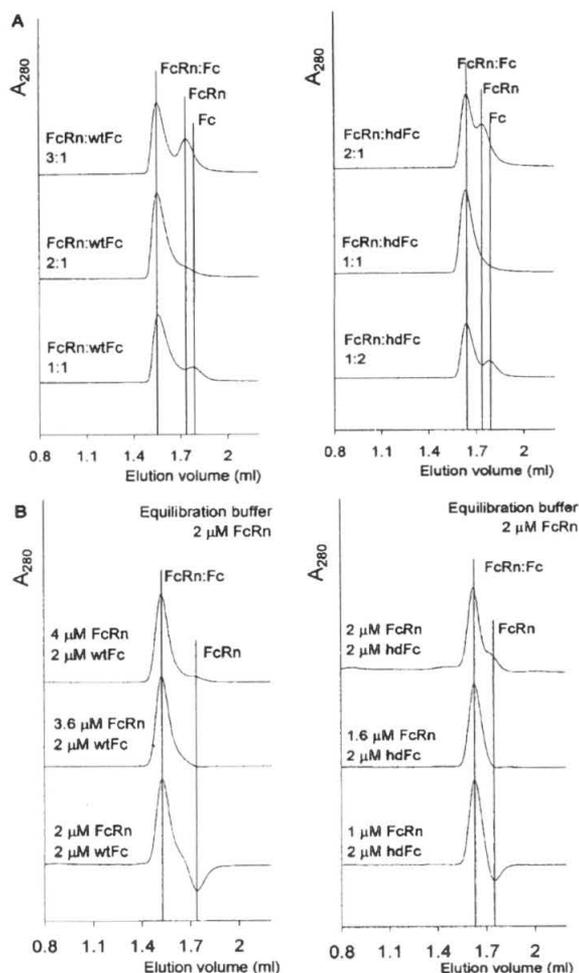
FIGURE 4: Gel filtration analyses of FcRn/Fc complexes. (A) Nonequilibrium gel filtration. FcRn and wtFc or hdFc were incubated at pH 6.0 at the indicated molar ratios and then passed over a gel filtration column run under nonequilibrium conditions to separate FcRn/Fc complexes from the free proteins. The identities of the proteins in each peak were confirmed by SDS−PAGE (data not shown). A single peak corresponding to an FcRn/Fc complex is formed at a 2:1 stoichiometry for FcRn/wtFc and a 1:1 stoichiometry for FcRn/hdFc. (B) Equilibrium gel filtration. FcRn was incubated with 2 μM wtFc or 2 μM hdFc at the indicated FcRn/Fc ratios in buffer containing 2 μM FcRn (equilibration buffer). Samples were injected onto a column equilibrated in the equilibration buffer. The peak that elutes first corresponds to an FcRn/Fc complex in equilibrium with free Fc; thus the concentration of bound Fc in the complex is lower than 2 μM. The second peak or trough is at the elution volume of free FcRn. Under equilibrium conditions, the stoichiometry of the solution complex can be determined from the chromatogram with a flat baseline by rounding up the ratio of injected proteins to the nearest integral value. Thus FcRn binds to wtFc with 2:1 stoichiometry and to hdFc with 1:1 stoichiometry under these conditions.

proteins chromatograph as a complex at a 1:1 molar ratio (Figure 4A).

Stoichiometries determined using conventional gel filtration are not definitive because the protein−protein complex is being assayed under nonequilibrium conditions. Thus if some of the complexes dissociate during the experiment, their

constituents can fail to rebind because they are being separated from each other during the chromatographic procedure. For example, high mannose carbohydrate-containing forms of mouse FcRn can form 1:1 complexes with Fc when assayed under nonequilibrium conditions (*10, 33*) but form 2:1 complexes when evaluated under equilibrium conditions (*10*). We therefore used the equilibrium gel filtration method of Hummel and Dreyer (*27*) to investigate the interactions of wtFc and hdFc with FcRn. As previously described (*10*), we equilibrated a gel filtration column with buffer containing a uniform concentration of FcRn. Pre-equilibrated complexes of FcRn plus Fc in different ratios were then injected over the gel filtration column. If the amount of additional FcRn injected with Fc is greater than or less than the amount required for formation of the FcRn/Fc complex, there is a peak (in the case of too much FcRn) or a trough (in the case of too little FcRn) at the position where free FcRn migrates. When the amount of additional FcRn injected with Fc is equal to the amount required for forming the FcRn/Fc complex, there is a flat baseline at the position where free FcRn migrates. Unless the protein concentrations greatly exceed the K_D of the interaction, however, the ratio of the concentration of added FcRn to the concentration of Fc will be a nonintegral value, from which the integral value corresponding to the stoichiometry generally can be derived by rounding up to the next integer. This is because the peak corresponding to the complex contains significant amounts of free Fc in equilbrium with the bound form, unless the experiment is conducted at concentrations that exceed the K_D by greater than 10-fold, which is usually not possible due to limiting amounts of purified protein.

We equilibrated a small (2.4 mL total volume) gel filtration column with 2 μM FcRn. Samples containing 2 μM wtFc or hdFc were incubated with various amounts of FcRn and chromatographed in the equilibration buffer containing FcRn. As shown in Figure 4B, a flat baseline is observed at the position where free FcRn migrates when 3.6 μM additional FcRn is injected with 2 μM wtFc; thus the stoichiometry of the FcRn/wtFc interaction is 2:1 under equilibrium conditions. By contrast, a flat baseline is observed when 1.6 μM hdFc is injected with 2 μM hdFc; thus at concentrations up to 2 μM, hdFc interacts with only one FcRn molecule (Figure 4B). The nbFc shows no specific interaction with FcRn (data not shown).

Comparison of wtFc and hdFc Binding to FcRn Using Surface Plasmon Resonance Assays. FcRn was covalently immobilized onto the surface of a biosensor chip, and binding of the Fc proteins was monitored in real time using a surface plasmon resonance-based binding assay, as previously described (*7, 8, 21, 31, 35*). We first analyzed the ability of nbFc to bind FcRn to check for residual binding at high concentrations. We found that nbFc does not generate a net binding response unless it is injected at concentrations exceeding 5 μM. At these concentrations, the responses are independent of concentration and are not reproducible from flow cell to flow cell or from chip to chip and thus represent nonspecific interactions with the biosensor chip (data not shown). For wtFc, the binding data were fit to a simple 1:1 interaction model and to a more complex model that assumes the response is due to two independent classes of noninteracting binding sites (heterogeneous ligand model) (Figure

5A). The heterogeneous ligand model fit the data better, as found in previous biosensor-based studies using intact IgG or Fc fragments (*31*). Using the heterogeneous ligand model, we derive a K_D of 1–6 nM for the high-affinity population, representing 55% of the binding sites, and a K_D of 143–218 nM for the low-affinity population, representing 45% of the binding sites. By contrast, the hdFc binding data fit the simple 1:1 interaction model quite well with a derived K_D of 75–96 nM (Figure 5A).

Because the hdFc/FcRn binding data fit the simple 1:1 interaction model, we inferred that the complex binding data observed in the wtFc/FcRn interaction result primarily from the presence of two potential FcRn binding sites on wtFc rather than from two populations of FcRn molecules on the chip surface. We therefore modeled the interaction of FcRn and wtFc as one side of wtFc binding the coupled FcRn followed by the wtFc/FcRn complex binding another FcRn using the second FcRn binding site on wtFc (bivalent analyte model). Using this model, a K_D is derived for each binding event, neither of which represents the affinity of wtFc being bound on both sides by FcRn; i.e., both K_D's represent distinct microscopic binding events. When the wtFc binding data are analyzed using this model, the K_D for the first binding event is in the range of 68–116 nM, and the second K_D is 317–487 nM (Figure 5A). Thus the K_D for the binding of FcRn to one site on wtFc is comparable to the 75–96 nM K_D derived for the binding of FcRn to hdFc. From these results, we conclude that the complex response exhibited by the interaction of FcRn with Fc or IgG on a biosensor chip is primarily the result of FcRn binding IgG or Fc at both FcRn binding sites.

We previously noted that the affinities of FcRn/IgG complexes are higher when FcRn, rather than IgG, is immobilized on the biosensor chip (*31*). All IgG subtypes tested showed a systematic coupling-dependent affinity difference, such that the high-affinity K_D when FcRn was immobilized (determined using a two-site heterogeneous ligand model) ranged between 15 and 93 nM, whereas the high-affinity K_D when IgG was immobilized ranged between 74 and 740 nM (*30*). To determine if hdFc exhibited a similar coupling-dependent affinity difference, we compared the binding of FcRn to wtFc and hdFc when each was immobilized to a biosensor chip. In both cases, the binding data could be fit to the simple 1:1 interaction model, and the derived K_D was 450–500 nM (Figure 5B). Thus both hdFc and wtFc bind to FcRn with lower affinity when they are immobilized than when they are injected over immobilized FcRn.

DISCUSSION

FcRn functions in the transport of IgG across epithelia and in the protection of IgG from catabolism in the serum (reviewed in refs *1–3*). Crystals of a complex between rat FcRn and the Fc fragment of IgG reveal an extended oligomeric ribbon of FcRn dimers bridged by homodimeric Fc's (*6*) (Figure 1). We have suggested that this ribbon forms under physiological conditions, such as the inside of an acidic transport vesicle, and that ribbon formation could serve as a component of a trafficking signal for directing vesicles containing bound IgG to their correct destination (*9*). To investigate this hypothesis, we have initiated a systematic
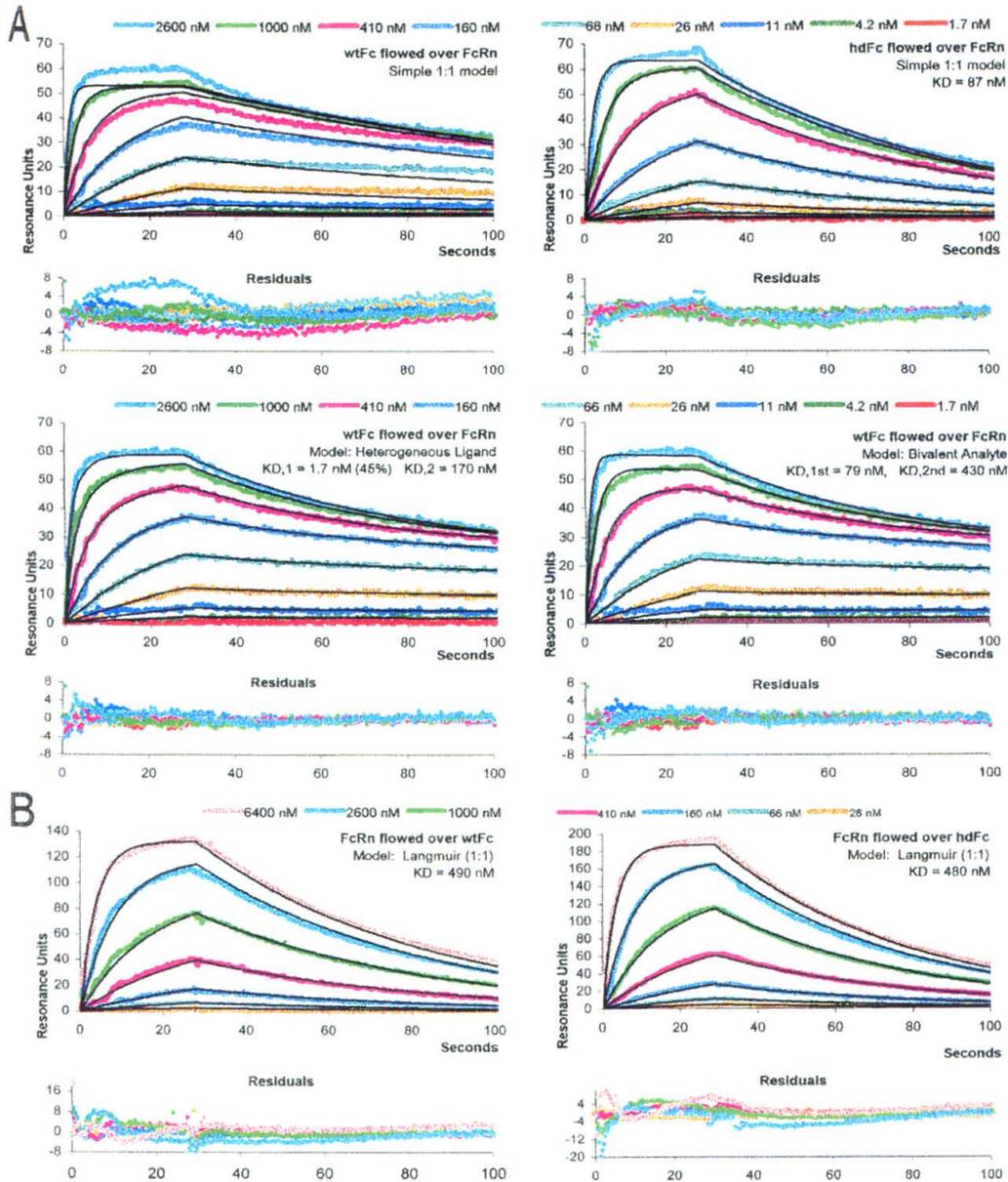
FIGURE 5: Biosensor analyses of FcRn/Fc complexes. FcRn or Fc was coupled to a biosensor chip at ~200, ~400, and ~1500 RU. Sensorgrams (colored lines) were analyzed using a kinetics-based method for the two lower coupling densities. An equilibrium-based method was used for analyses of data from all three coupling densities, yielding results comparable to those of the kinetics-based analyses (data not shown). In each panel, the model used to fit the data (thin black lines overlaid with the observed response) is listed along with derived affinity constant(s). The simple 1:1 model fits the data to the reaction FcRn + Fc $\rightleftharpoons$ FcRn:Fc. The heterogeneous ligand model assumes that there are two populations of FcRn on the chip and fits the data according to the following reactions: FcRn + Fc $\rightleftharpoons$ FcRn:Fc and FcRn* + Fc $\rightleftharpoons$ FcRn*:Fc. K_D's ($K_{D,1}$ or $K_{D,2}$) and the percentage of the total response due to each population of FcRn are derived for each reaction. The bivalent analyte model fits the data according to the following sequential reactions: FcRn + Fc $\rightleftharpoons$ FcRn:Fc and FcRn: Fc + FcRn $\rightleftharpoons$ FcRn:Fc:FcRn. K_D's ($K_{D,first}$ and $K_{D,second}$) are derived for each reaction. (A) Sensorgrams from kinetics-based experiments in which the indicated Fc is flowed over FcRn. One representative set of injections from experiments performed in duplicate or triplicate is shown for each interaction on a chip coupled to ~400 RU. Similar results were obtained for the ~200 RU coupling density chips (wtFc: $K_{D,1} = 4$ nM (58%); $K_{D,2} = 185$ nM (42%); $K_{D,first} = 101$ nM; $K_{D,second} = 336$ nM. hdFc: $K_D = 86$ nM). (B) Sensorgrams from kinetics-based experiments in which FcRn is flowed over the indicated Fc. One representative set of injections from experiments performed in duplicate or triplicate is shown for each interaction on a chip coupled to ~200 RU. Similar results were obtained for the ~400 RU coupling density chips (wtFc: $K_D = 500$ nM. hdFc: $K_D = 460$ nM).

characterization of the interaction between FcRn and its Fc ligand.

Many biochemical studies of FcRn, a membrane-bound receptor that normally interacts with its ligand in the lumen of intracellular transport vesicles, have been done in solution using a soluble version of the receptor. Under equilibrium conditions in solution, the FcRn/ligand complex is composed of three molecules, two FcRn's and one Fc (*10, 11*). In the FcRn/Fc cocrystals, there are two distinct possibilities that could account for the 2:1 complex observed in solution. In one, a dimer of FcRn molecules binds to only one FcRn binding site on homodimeric Fc. In the other, single FcRn molecules bind to both FcRn binding sites on Fc (*6*).

Here we describe the use of recombinant Fc proteins containing zero, one, or two binding sites for FcRn (nbFc, hdFc, and wtFc, respectively) to determine the nature of the FcRn/Fc complex formed in solution. The experimental results consistently demonstrate that the 2:1 FcRn/Fc solution complex consists of two FcRn molecules binding to both sides of wtFc. First, using an assay involving FcRn bound to a solid support, we show that more than one FcRn molecule can bind to wtFc but not to hdFc (Figure 3). In addition, gel filtration analyses under both equilibrium and nonequilibrium conditions demonstrate that hdFc forms a 1:1 complex with FcRn, whereas wtFc forms a 2:1 FcRn/Fc complex (Figure 4). The 1:1 nature of the FcRn complex with hdFc is reflected in biosensor analyses of this interaction, in that binding data involving hdFc could be fit to a simple 1:1 binding model. By contrast, binding data for the interaction of coupled FcRn with wtFc or intact IgG must be fit to more complex binding models that incorporate binding to the second FcRn binding site on Fc (Figure 5A). Combined with a previous demonstration that soluble FcRn is monomeric at micromolar concentrations in solution (*22*), the present results establish that FcRn does not dimerize in solution, either alone or when bound to Fc.

The result that the 2:1 FcRn/Fc complex formed in solution does not include FcRn dimers cannot be used to infer that receptor dimerization does not occur under physiological conditions. In biochemical experiments such as those described here, soluble FcRn is studied at relatively low protein concentrations (micromolar) that would not be expected to favor formation of receptor dimers or the oligomeric ribbon. In vivo, however, receptors are tethered to a membrane under conditions of high effective molarity in which receptor dimerization and oligomeric ribbon formation could be facilitated. Tethering of soluble FcRn to a biosensor chip may to some extent mimic the high local protein concentrations found in a membrane by facilitating dimerization of FcRn. Previous studies suggest that FcRn can dimerize on a biosensor chip, in that mutations at the FcRn dimer interface that do not directly contact IgG resulted in reduced affinities for IgG (*8*). In addition, the previous observation that the affinity between FcRn and Fc or IgG is highest when the receptor rather than the ligand is coupled to a biosensor chip (*31*) can now be interpreted by assuming that FcRn can dimerize when coupled to a biosensor chip. This orientation effect is not due to the ability of FcRn to bind IgG and wtFc at two sites because it is also produced by hdFc, which can only be bound on one side by FcRn. A reasonable explanation for this effect is that FcRn can form dimers when coupled to the surface of the biosensor chip at high effective molarity and that these dimers bind Fc and IgG more stably than does monomeric FcRn, which is the predominant species binding to immobilized Fc or IgG.

The use of Fc molecules that contain zero, one, or two FcRn binding sites has allowed the identification of the trimolecular 2:1 FcRn/Fc complex that forms under micromolar conditions in solution. Future studies of the interaction of these Fc molecules with membrane-bound FcRn will facilitate understanding of the more complex interactions between this receptor and its ligand under physiological conditions.

ACKNOWLEDGMENT

REFERENCES

1. Junghans, R. P. (1997) *Immunol. Res. 16*, 29−57.
2. Ghetie, V., and Ward, E. S. (1997) *Immunol. Today 18*, 592−598.
3. Simister, N. E., Israel, E. J., Ahouse, J. C., and Story, C. M. (1997) *Biochem. Soc. Trans. 25*, 481−486.
4. Simister, N. E., and Mostov, K. E. (1989) *Nature 337*, 184−187.
5. Burmeister, W. P., Gastinel, L. N., Simister, N. E., Blum, M. L., and Bjorkman, P. J. (1994) *Nature 372*, 336−343.
6. Burmeister, W. P., Huber, A. H., and Bjorkman, P. J. (1994) *Nature 372*, 379−383.
7. Raghavan, M., Chen, M. Y., Gastinel, L. N., and Bjorkman, P. J. (1994) *Immunity 1*, 303−315.
8. Vaughn, D. E., Milburn, C. M., Penny, D. M., Martin, W. L., Johnson, J. L., and Bjorkman, P. J. (1997) *J. Mol. Biol. 274*, 597−607.
9. Raghavan, M., and Bjorkman, P. J. (1996) *Annu. Rev. Cell Biol. 12*, 181−220.
10. Sánchez, L. M., Penny, D. M., and Bjorkman, P. J. (1999) *Biochemistry 38*, 9471−9476.
11. Huber, A. H., Kelley, R. F., Gastinel, L. N., and Bjorkman, P. J. (1993) *J. Mol. Biol. 230*, 1077−1083.
12. Weng, Z., Gulukota, K., Vaughn, D. E., Bjorkman, P. J., and DeLisi, C. (1997) *J. Mol. Biol. 282*, 217−225.
13. Kim, J.-K., Tsen, M.-F., Ghetie, V., and Ward, E. S. (1994) *Eur. J. Immunol. 24*, 2429−2434.
14. Kim, J.-K., Tsen, M.-F., Ghetie, V., and Ward, E. S. (1994) *Scand. J. Immunol. 40*, 457−465.
15. Kim, J. K., Tsen, M. F., Ghetie, V., and Ward, E. S. (1994) *Eur. J. Immunol. 24*, 542−548.
16. Ghetie, V., Popov, S., Borvak, J., Radu, C., Matesoi, D., Medesan, C., Ober, R. J., and Ward, E. S. (1997) *Nat. Biotechnol. 15*, 637−640.
17. Medesan, C., Matesoi, D., Radu, C., Ghetie, V., and Ward, E. S. (1997) *J. Immunol. 158*, 2211−2217.
18. Medesan, C., Cianga, P., Mummert, M., Stanescu, D., Ghetie, V., and Ward, E. S. (1998) *Eur. J. Immunol. 28*, 2092−2100.
19. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
20. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C. (1991) *Sequences of proteins of immunological interest*, U.S. Department of Health and Human Services, Bethesda, MD.
21. Raghavan, M., Bonagura, V. R., Morrison, S. L., and Bjorkman, P. J. (1995) *Biochemistry 34*, 14649−14657.
22. Gastinel, L. N., Simister, N. E., and Bjorkman, P. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 638−642.
23. Bebbington, C. R., and Hentschel, C. C. G. (1987) The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells, in *DNA Cloning: A*

Practical Approach (Glover, D. M., Ed.) pp 163−188, IRL Press, Oxford.

24. Chapman, T. L., and Bjorkman, P. J. (1997) *J. Virol.* (in press).

25. GCG, W. (1994) *Program manual for the Wisconsin Package*, Genetics Computer Group, Madison, WI.

26. Gill, S. C., and Von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.

27. Hummel, J. P., and Dreyer, W. J. (1962) *Biochim. Biophys. Acta 63*, 530−532.

28. Fägerstam, L. G., Frostell-Karlsson, A., Karlsson, R., Persson, B., and Rönnber, I. (1992) *J. Chromatogr. 597*, 397−410.

29. Malmqvist, M. (1993) *Nature 361*, 186−187.

30. Karlsson, R., and Fält, A. (1997) *J. Immunol. Methods 200*, 121−133.

31. Vaughn, D. E., and Bjorkman, P. J. (1997) *Biochemistry 36*, 9374−9380.

32. Ghetie, V., Hubbard, J. G., Kim, J.-K., Tsen, M.-F., Lee, Y., and Ward, E. S. (1996) *Eur. J. Immunol. 26*, 690−696.

33. Popov, S., Hubbard, J. G., Kim, J.-K., Ober, B., Ghetie, V., and Ward, E. S. (1996) *Mol. Immunol. 33*, 521−530.

34. Rodewald, R., and Kraehenbuhl, J.-P. (1984) *J. Cell Biol. 99*, S159−S164.

35. Raghavan, M., and Bjorkman, P. J. (1995) *Structure 3*, 331−333.

Chapter 3:

Crystal structure at 2.8 Å of an FcRn/Heterodimeric Fc complex: Mechanism of pH-dependent binding

This chapter describes the solution of FcRn/heterodimeric Fc complex crystals as well as the structure of the non-binding Fc. Dr. Anthony West assisted in the solution of FcRn/hdFc structure. Lu Gan purified the non-binding Fc, generated the non-binding Fc crystals, collected the diffraction data, and participated in the solution and refinement of the non-binding Fc.

Crystal Structure at 2.8 Å of an FcRn/Heterodimeric Fc Complex: Mechanism of pH-Dependent Binding

W. Lance Martin,* Anthony P. West, Jr.,* Lu Gan,*
and Pamela J. Bjorkman*†‡
*Division of Biology
†Howard Hughes Medical Institute
California Institute of Technology
Pasadena, California 91125

Summary

The neonatal Fc receptor (FcRn) transports immunoglobulin G (IgG) across epithelia, binding IgG in acidic vesicles (pH $\leq$ 6.5) and releasing IgG in the blood at pH 7.4. Well-ordered FcRn/Fc crystals are prevented by the formation of "oligomeric ribbons" of FcRn dimers bridged by Fc homodimers, thus we crystallized a 1:1 complex between rat FcRn and a heterodimeric Fc containing only one FcRn binding site. The 2.8 Å complex structure demonstrates that FcRn uses its $\alpha 2$ and $\beta 2$-microglobulin domains and carbohydrate to interact with the Fc $C_\gamma 2$–$C_\gamma 3$ interface. The structure reveals conformational changes in Fc and three titratable salt bridges that confer pH-dependent binding, and can be used to guide rational design of therapeutic IgGs with longer serum half-lives.

Introduction

FcRn mediates transport of maternal IgG across the neonatal intestine in rodents and the placenta in humans, thereby conferring humoral immunity to the fetus or newborn against antigens encountered by the mother. In addition, FcRn protects IgG from degradation by serving as the receptor responsible for maintenance of the long half-life and high concentrations of IgG in serum (Simister and Mostov, 1989; Ghetie and Ward, 2000). In both its transport and protection receptor roles, FcRn binds IgG with nanomolar affinity at acidic pH ($\leq$6.5) in intracellular transport vesicles and releases IgG upon encountering the basic pH of the bloodstream (7.4) (Ghetie and Ward, 2000). Understanding the details of the FcRn interaction with IgG is critical for efforts to increase the serum half-lives of antibody-based drugs and to deliver therapeutic IgGs across the placenta.

FcRn shares 22%–29% sequence identity with class I major histocompatibility (MHC) molecules (Simister and Mostov, 1989), which present peptide antigens to cytotoxic T cells. Both types of proteins are heterodimers composed of the soluble light chain $\beta 2$-microglobulin ($\beta 2m$) and a membrane-bound heavy chain that includes three extracellular domains ($\alpha 1$, $\alpha 2$, and $\alpha 3$), a single-pass transmembrane region, and a short cytoplasmic domain. Crystal structures of rat (Burmeister et al., 1994a) and human (West and Bjorkman, 2000) FcRn confirmed the structural similarity with class I MHC molecules, except that FcRn has a narrowed and nonfunc-

tional version of the class I MHC peptide binding groove. The low resolution crystal structure (6.5 Å) of a rat FcRn/Fc complex revealed that the side of an FcRn $\alpha 1$–$\alpha 2$ domain platform interacts with the $C_\gamma 2$–$C_\gamma 3$ domain interface on each chain of the Fc homodimer (Burmeister et al., 1994b). Many different surfaces on class I MHC and class I homologs are employed for interactions with protein ligands (Strong, 2000). By contrast, the Fc $C_\gamma 2$–$C_\gamma 3$ interface is the binding site for a number of proteins that bind IgG, including protein A (Deisenhofer, 1981), protein G (Sauer-Eriksson et al., 1995), rheumatoid factor (Corper et al., 1997), peptides selected for high-affinity Fc binding (DeLano et al., 2000), and the herpes simplex virus IgG binding protein gE-gl (Chapman et al., 1999). The limited resolution of the FcRn/Fc cocrystals prohibited detailed comparisons with the other Fc binding proteins and analyses of potential conformational changes induced by binding. Growth of well-ordered cocrystals is apparently prevented by the packing, in which Fc homodimers bridge between dimers of FcRn heterodimers to create an "oligomeric ribbon" (Figure 1A).

In order to obtain an FcRn/Fc cocrystal with a different packing arrangement, we engineered a heterodimeric version of Fc (hdFc) that cannot bridge between FcRn molecules because it contains only a single FcRn binding site (Martin and Bjorkman, 1999). The hdFc is composed of a wild-type rat IgG2a Fc region (wtFc) covalently linked via hinge region disulfide bonds to a mutant rat IgG2a Fc (nonbinding Fc; nbFc) with substitutions that disrupt FcRn binding (Thr-252 to Gly, Ile-253 to Gly, Thr-254 to Gly, His-310 to Glu, His-433 to Glu, and His-435 to Glu). In solution, hdFc forms a 1:1 complex with FcRn, in contrast to wtFc homodimers, which form 2:1 FcRn/Fc complexes (Martin and Bjorkman, 1999). Here we describe structures of a 1:1 FcRn/hdFc complex and a nbFc homodimer solved to 2.8 Å and 2.7 Å, respectively. The FcRn/hdFc structure reveals the molecular mechanism for the pH dependence of the FcRn/IgG interaction and shows that ordered carbohydrate from FcRn participates in binding Fc. Comparisons of the structures of free and FcRn-bound Fc molecules reveal domain rearrangements distant from the FcRn binding site and asymmetry in Fc that may result in negative cooperativity for binding the second FcRn to Fc.

Results

FcRn/hdFc and nbFc Structures

Secreted forms of rat FcRn, hdFc, and nbFc were expressed in CHO cells and purified as described (Martin and Bjorkman, 1999). The FcRn/hdFc and nbFc structures were solved at pH 5.4 (complex crystals) or pH 6.4 (nbFc crystals) by molecular replacement using the structures of rat FcRn (Burmeister et al., 1994a; Vaughn and Bjorkman, 1998) and/or human Fc (Deisenhofer, 1981) (Table 1). The FcRn/hdFc structure consists of a 1:1 complex in which FcRn interacts specifically with the wtFc (proximal) chain of the hdFc with the side of

‡To whom correspondence should be addressed (e-mail: bjorkman@its.caltech.edu).
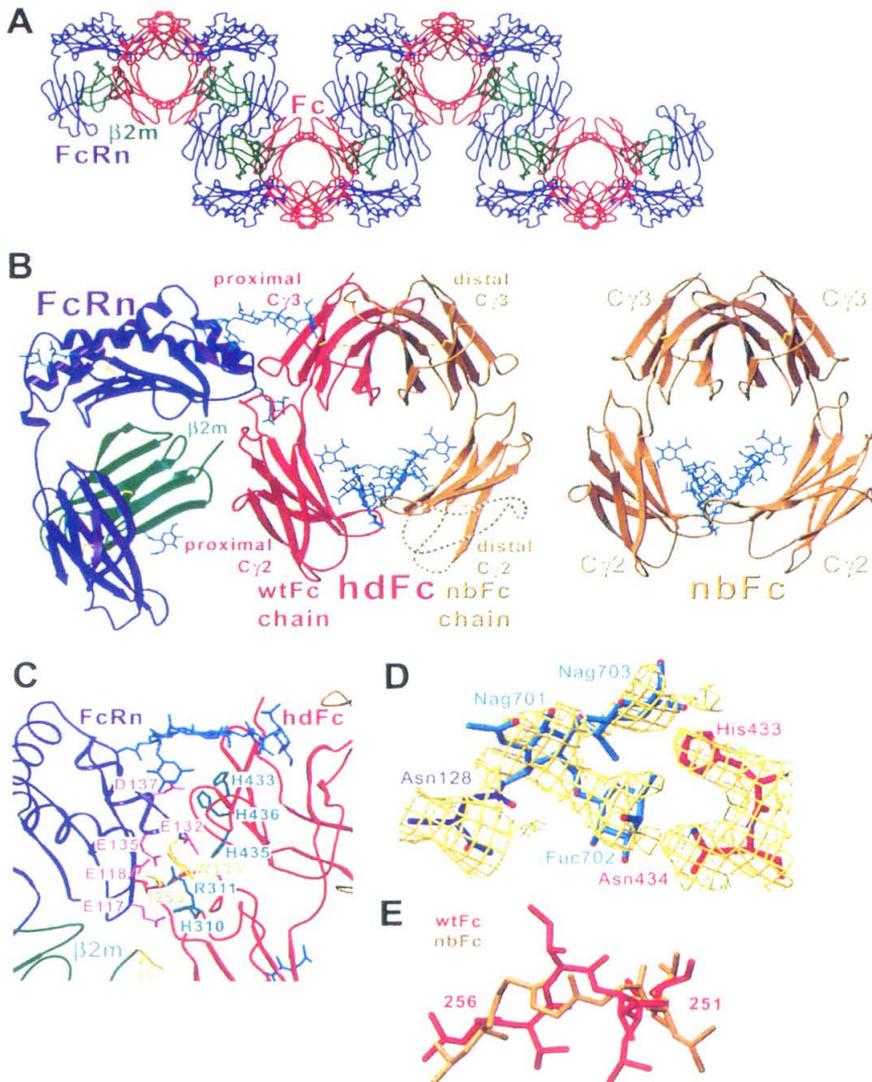
Figure 1. FcRn/Fc, FcRn/hdFc, and nbFc Structures

(A) FcRn/Fc complexes in the oligomeric ribbon observed in crystals of FcRn bound to wtFc. FcRn/Fc crystals grown using human, mouse, or rat FcRn and human, mouse, or rat Fc subclasses all appear to contain the oligomeric ribbon packing in which FcRn dimers are bridged by Fc homodimers. Such crystals diffract aniostropically to 3.5 Å–8 Å, with the highest resolution in the direction of the long axis of the FcRn dimer.

(B) Ribbon diagrams of the structures of FcRn/hdFc and nbFc. Ordered N-linked carbohydrates are shown in ball-and-stick representation. Disulfide bonds are yellow. Regions of disorder in the distal $C_\gamma2$ domain are shown as dashed lines. The FcRn/hdFc complexes are packed in the crystals such that the nbFc chain of the hdFc contacts an FcRn in an adjacent FcRn/hdFc complex. This interaction involves a face of the FcRn $\alpha3$ domain opposite from the Fc binding site, and the buried surface area (577 Å² total) is near the mean size buried in typical crystal contacts (570 Å²) (Janin, 1997), thus it is a nonspecific interaction.

(C) Close-up of the FcRn/hdFc interface. Interface residues are turquoise (positively charged), pink (negatively charged), and yellow (hydrophobic). The carbohydrate attached to residue Asn-87 was omitted for clarity.

(D) The FcRn/hdFc model in the region of the N-linked carbohydrate attached to FcRn Asn-128 superimposed on a 2.8 Å SIGMAA-weighted $2F_o$-F_c annealed omit electron density map contoured at 1.0 σ.

(E) Comparison of the Fc 251 to 256 loop in the wt (red) and nb (gold) sides of hdFc (Cα rms deviation of 1.78 Å).

Table 1. Data Collection and Refinement Statistics for the FcRn/hdFc and nbFc Crystal Structures

	FcRn/hdFc	nbFc
Data Collection		
Space group	P2₁2₁2₁	P2₁
Unit cell (Å)	68.4, 74.4, 196.6	42.5, 73.4, 82.2, β = 103°
Resolution (Å)	20–2.8 (2.9–2.8)	30–2.6 (2.7–2.6)
Wavelength (Å)	1.00	1.54
Mosaicity (°)	0.7	0.6
Observations	96,527 (7836)	35,621 (3003)
Unique reflections	25,112 (2234)	14,919 (1323)
Completeness (%)[a]	98.7 (89.8)	97.5 (87.2)
I/σI	18.9 (3.0)	13.4 (3.6)
R_{merge} (%)[b]	8.0 (30.0)	6.8 (26.1)
Refinement		
Resolution (Å)	20–2.8	20–2.7
Work reflections	23,810	12,096
Test reflections	1195	1375
R_{cryst} (%)[c]	22.2	24.1
R_{free}[d]	29.3	27.9
No. of protein atoms	6241	3264
No. of waters	6	0
No. of carbohydrate atoms	323	198
Average B factor (Å²)	79.1	41.2
Anisotropic B correction	B11 = −24.12	B11 = 7.9 B12 = 0.00
	B22 = 5.95	B22 = −5.2 B13 = −0.66
	B33 = 18.16	B33 = −2.7 B23 = 0.00
Rms deviations from ideal		
Bond lengths (Å)	0.009	0.008
Bond angles (°)	1.6	1.4
Ramachandran plot quality		
Most favored (%)	77.8	89.1
Additionally favored (%)	20.8	9.8
Generously allowed (%)	1.4	1.1
Disallowed (%)	0.0	0.0

Statistics in parentheses refer to the highest resolution bin.

[a] Completeness = (number of independent reflections)/total theoretical number.

[b] R_{merge} (I) = $(\Sigma|I(i) - <I(h)>|/\Sigma I(i))$, where I(i) is the ith observation of the intensity of the hkl reflection and <I> is the mean intensity from multiple measurements of the h,k,l reflection.

[c] R_{cryst} (F) = $\Sigma h||Fobs(h)| - |Fcalc(h)||/\Sigma h|Fobs(h)|$, where |Fobs(h)| and |Fcalc(h)| are the observed and calculated structure factor amplitudes for the h,k,l reflection.

[d] R_{free} is calculated over reflections in a test set not included in atomic refinement.

its α1–α2 domain platform recognizing the Fc $C_\gamma 2$–$C_\gamma 3$ domain interface (Figures 1B and 1C). The oligomeric ribbon is completely disrupted, since FcRn does not dimerize and hdFc does not bridge between FcRn molecules in the FcRn/hdFc crystals. There is ordered electron density for FcRn, the FcRn/Fc interface, N-linked carbohydrates on FcRn (Figure 1D) and Fc, the proximal (FcRn-bound) side of the hdFc excluding the B-C loop of the $C_\gamma 2$ domain, and the $C_\gamma 3$ domain of the distal (nbFc) chain. In the distal $C_\gamma 2$ domain, only the $C_\gamma 3$-nearest half, the intradomain disulfide, and the N-linked carbohydrates show ordered electron density, thus the remainder of the $C_\gamma 2$ domain was modeled using coordinates from the nbFc structure. Disorder of the same portion of the distal $C_\gamma 2$ domain is also observed in a 5 Å structure derived from FcRn/hdFc crystals collected at room temperature (W. L. M. and P. J. B., unpublished results). The same portions of the $C_\gamma 2$ domains of human Fc (Deisenhofer, 1981), a human Fc/protein A fragment complex (Deisenhofer, 1981), and an intact antibody (Harris et al., 1997) also show disorder.

The structure of the nbFc homodimer (Figure 1B) closely resembles structures of human Fc (Deisenhofer, 1981; Sauer-Eriksson et al., 1995; Corper et al., 1997; Harris et al., 1997; DeLano et al., 2000). The most notable difference concerns the loop containing residues 252–254, which are replaced by glycines. This loop is rearranged relative to the corresponding loops in human Fc or the wtFc chain of hdFc (Figure 1E), suggesting that movement of the 251–256 loop as well as the absence of FcRn-interacting side chains contributes to the lack of binding between nbFc and FcRn.

FcRn/Fc Interface

The FcRn/Fc binding interface spans a large surface area and is highly complementary. The buried surface area at the interface (1870 Å²) (Table 2) is slightly larger than average protein–protein recognition interfaces (1560–1700 Å²) (Jones and Thornton, 1996) and larger than areas buried at the interface between Fc and other proteins that bind to $C_\gamma 2$–$C_\gamma 3$ interdomain region of Fc (protein A, protein G, and rheumatoid factor) (Table 2). In addition, the shape correlation statistic (S_c) (Lawrence and Colman, 1993) for the FcRn/hdFc complex is higher than indices for these and other protein/Fc complexes (Table 2). On FcRn, the Fc binding site encompasses

Table 2. Characteristics of the FcRn/hdFc and Other Protein–Protein Interfaces

Protein(s)	Interface Area (Å²)	Shape Complementarity	Number of Salt Bridges	Percentage Nonpolar Interface Area	Percentage Polar Interface Area	Percentage Charged Interface Area	Proximal, Distal Fc C_γ–C_γ3 Angles
FcRn/hdFc	1870	0.71	4	56	28	16	89°, 81°
Protein A/Fc	1320	0.66	0	61	36	3	99°, 99°
Protein G/Fc	1370	0.56	2	56	33	11	100°, 100°
RF/Fc	1490	0.69	2	47	34	9	93°, 93°
FcγRIII/Fc	1730	0.67	0	67	24	9	93°, 106°
FcϵRI/Fcϵ	1460	0.69	2	62	27	11	85°, 98°
Free hFc							98°, 95°
Free nbFc							92°, 92°
Intact mIgG							98°, 94°

RF denotes rheumatoid factor. Interface area in column 2 is the total buried surface area at a protein–protein interface and was calculated as described in Experimental Procedures. Shape correlation statistics (S_c) were calculated using SC (Lawrence and Colman, 1993) ($S_c = 1$ for interfaces with perfect fits). Salt bridges are defined as atoms on each interacting protein with opposite charges that are within 3.5 Å of each other. Interface buried surface area in columns 5–7 is defined and characterized as nonpolar, polar, or charged according to the criteria of Lo Conte et al. (1999). The percentage in each category is calculated for all protein atoms at the interface. The interdomain angle is the angle between the long axes of the C_γ2–C_γ3 domains of Fc calculated as described (Su et al., 1998).

the C-terminal portion of the α2 domain α helix and the first residues of β2m. On Fc, the FcRn binding site encompasses the residues identified by DeLano et al. (2000) as common to all interactions involving the Fc C_γ2–C_γ3 interface (residues 252–254 in the C_γ2 A-B loop and residues 434–436 in strand G of the C_γ3 domain). In addition, the FcRn footprint on Fc includes residues in the C_γ2 E-F loop (309–311) (Table 3).

The center of the FcRn/Fc interface includes a core of hydrophobic residues in which FcRn Trp-133 contacts Fc Ile-253 and β2m Ile-1 from FcRn contacts Fc Pro-307 (Figure 1C). Substitution of FcRn Trp-133, β2m Ile-1, or Fc Ile-253 greatly reduces the interaction affinity (Table 3) (Popov et al., 1996; Vaughn et al., 1997; Shields et al., 2001). Surrounding the hydrophobic core containing Fc Ile-253 is a network of salt bridges involving FcRn residues Glu-117, Glu-118, Glu-132, and Asp-137 and Fc residues His-310, Arg-311, His-435, and His-436 (Figure 2A). Mutation of Glu-117, Glu-132, and Asp-137 on FcRn and His-310, Arg-311, His-435, and His-436 on Fc, either individually or in conjunction with other residues, lowers the binding affinity substantially (Table 3) (Kim et al., 1994a, 1999; Raghavan et al., 1995; Popov et al., 1996; Ghetie and Ward, 1997; Medesan et al., 1997, 1998; Vaughn et al., 1997; Shields et al., 2001).

N-linked Carbohydrate Attached to FcRn Contributes to Fc Binding

Unanticipated from the low-resolution FcRn/Fc structure (Burmeister et al., 1994b), we find extensive interactions involving N-linked carbohydrates from FcRn that contact Fc. Contacts between Fc and three sugar residues attached to FcRn Asn-128 account for 10%–15% of the buried surface area in the FcRn/Fc interface (Figures 1C and 1D; Table 3). The sugars contact four Fc residues, including His-433 (Figure 1D), previously suggested to contribute to binding of rat FcRn (Raghavan et al., 1995). The N-linked glycosylation site at Asn-128 is found in rodent (Simister and Mostov, 1989; Ahouse et al., 1993) but not human (Story et al., 1994) or bovine (Kacskovics et al., 2000), forms of FcRn (Table 3). In the case of mouse FcRn, differential glycosylation affects the receptor/ligand stoichiometry such that 1:1 FcRn/Fc

complexes can be isolated using high-mannose forms of FcRn (Popov et al., 1996; Sánchez et al., 1999), whereas 2:1 complexes are observed under the same conditions using FcRn with complex carbohydrates (Sánchez et al., 1999). Contact between Fc and complex carbohydrate–specific sugar residues on FcRn (fucose and N-acetyl-glucosamine) (Table 3) suggests that maximal Fc binding affinity requires complex carbohydrate attached to FcRn Asn-128, thus the affinity for binding a second FcRn to Fc should be lower in high-mannose-containing rather than complex carbohydrate–containing forms of rodent FcRn (Schuck et al., 1999).

Conformational Changes at the FcRn/hdFc Interface

FcRn does not undergo a major conformational change upon binding Fc, but there are slight but significant rearrangements in FcRn side chains at the binding interface. In complexed FcRn, Glu-135 adopts a different rotamer to form hydrogen bonds with backbone amide nitrogens in Fc residues 253 and 254 (Figure 2B). β2m Ile-1 swings out to make van der Waals contacts with Fc residues Val-309 and Pro-307 (Figure 2B). Another interface residue, FcRn Glu-132, adopts a different rotamer to form a salt bridge with Fc His-435 (Figures 2A and 2B). Overall, however, the structures of free and Fc-bound FcRn superimpose well (0.89 Å rms deviation for all Cα atoms).

Fc side chains at the FcRn binding site are also reoriented in apparent response to receptor binding. His-436, Arg-311, and Ile-253 assume different rotamers in the structure of FcRn-bound Fc compared with structures of free Fc or nbFc (Figure 2B). All three differences likely facilitate FcRn binding to Fc: His-436 forms a salt bridge with FcRn Asp-137, Arg-311 forms a salt bridge with FcRn Glu-118 and binds an ordered water also bound by Glu-117, and Ile-253 makes van der Waals contacts with FcRn Trp-133.

In addition to side chain changes at the FcRn binding site, we find quaternary structure rearrangements distant from the FcRn binding site when the structures of free Fc (human Fc [Deisenhofer, 1981] or the nbFc homodimer), FcRn-bound Fc, and human Fc bound to other proteins (Deisenhofer, 1981; Sauer-Eriksson et al.,

Table 3. Interacting Residues at the FcRn/hdFc Interface

FcRn Residue (ASA)	RMHB	Mutation (Effect on Affinity)	Fc Residue (ASA)	R 1222 abc	M 1222 abc	H 1234	B 123 a	Mutation (Effect of Affinity for mFcRn or rFcRn)	Mutation (Relative Binding to hFcRn)
86 (1.3%)	IL--	N84GQ85YI86Y (2× ↑)	254 (5.5%)	TTST	TTSS	SSSS	STS	M252GI253GS254G (90× ↓)	S254A (<0.10)
90 (0.5%)	FYYF	—	254 (5.5%)	TTST	TSSS	SSSS	STS	M252GI253GS254G (90× ↓)	S254A (<0.10)
117 (4.6%)	EEEE	E117S (»300× ↓)	309 (2.8%)	LVQQ	MQQQ	LVLI	QQQ	L309GH310GR311G (80× ↓)	—
117 (4.6%)	EEEE	E117S (»300× ↓)	310 (1.3%)	HHHH	HHHH	HHHH	HHH	H310A (8-33× ↓)	—
118 (1.8%)	EEEE	—	311 (4.5%)	QRQQ	QQQQ	QQQQ	QQQ	L309GH310GR311G (80× ↓)	Q311A (1.62)
119 (0.4%)	FFFF	—	253 (7.0%)	IIII	IIII	IIII	III	M252GI253GS254G (90× ↓)	I253A (<0.10)
132 (3.4%)	EEDD	E132QE135Q (»300× ↓)	435 (2.9%)	HHHH	HHYH	HHRH	HHH	H435A (13-20× ↓)	H435A (<0.10)
133 (6.8%)	WWWW	W132A (»300× ↓)	253 (7.0%)	IIII	IIII	IIII	III	I253A (3000× ↓)	I253A (<0.10)
133 (6.8%)	WWWW	W133A (»300× ↓)	310 (1.3%)	HHHH	HHHH	HHHH	HHH	H310A (8-33× ↓)	—
133 (6.8%)	WWWW	W133A (»300× ↓)	311 (4.5%)	QRQQ	QQQQ	QQQQ	QQQ	L309GH310GR311G (80× ↓)	Q311A (1.62)
133 (6.8%)	WWWW	W133A (»300× ↓)	314 (0.8%)	LLMM	LLMM	LLMM	TTL	—	—
133 (6.8%)	WWWW	W133A (»300× ↓)	435 (2.9%)	HHHH	HHYH	HHRH	HHH	H435A (13-20× ↓)	H435A (<0.10)
135 (2.3%)	EEEE	E132QE135Q (»300× ↓)	252 (1.1%)	TTLM	TMMM	MMMM	TMT	M252GI253GS254G (90× ↓)	M252A (1.0)
135 (2.3%)	EEEE	E132QE135Q (»300× ↓)	253 (7.0%)	IIII	IIII	IIII	III	I253A (3000× ↓)	I253A (<0.10)
135 (2.3%)	EEEE	E132QE135Q (»300× ↓)	254 (5.5%)	TTST	TTSS	SSSS	STS	M252GI253GS254G (90× ↓)	S254A (<0.10)
136 (0.0%)	TTAS	—	434 (5.9%)	NNNN	NNNN	NNNN	NNN	N434A (2× ↓ to no effect)	N434A (3.46)
137 (1.8%)	DELR	D137N (»300× ↓)	434 (5.9%)	NNNN	NNNN	NNNN	NNN	N434A (2× ↓ to no effect)	N434A (3.46)
137 (1.8%)	DELR	D137N (»300× ↓)	436 (1.9%)	HHHH	HHYH / L	YYFY / Y	YYY	H436A (3-6× ↓)	Y436A (<0.10)
β2m 1 (5.1%)	IIII	I1A (»40× ↓)	307 (1.5%)	PPPI	PPPP	TTTT	RPR	—	—
β2m 2 (1.6%)	QQQQ	Q2A (2× ↑)	288 (0.8%)	QQQQ	QQQQ	KKKK	TRR	—	K288A (0.38)
FcRn N128 (sugar)									
FUC 702 (2.4%)	+ + - -	—	433 (3.1%)	HHHH	HHKH	HHHH	HHR	H433A (2× ↓ to no effect)	H433A (0.41)
FUC 702 (2.4%)	+ + - -	—	434 (5.9%)	NNNN	NNNN	NNNN	NNN	N434A (2× ↓ to no effect)	N434A (3.46)
FUC 709 (8.2%)	+ + - -	—	348 (0.8%)	VVVV	VVVV	VVVV	VVV	—	—
FUC 709 (8.2%)	+ + - -	—	439 (2.9%)	KKKK	KKKK	KKKK	KKK	—	K439A (1.0)

Pairwise interactions were identified by contact analysis in CNS (Brünger et al., 1998) using the FcRn/hdFc structure (rat FcRn bound to a rat IgG2a Fc). Contacting residues were defined as those containing an atom within 4.0 Å of any nonhydrogen atom on the partner molecule. Counterpart residues in mouse (M) (Ahouse et al., 1993), human (H) (Story et al., 1994), and bovine (B) (Kacskovics et al., 2000) FcRn and Fc regions are listed, with Fc sequences divided into subclasses (Kabat et al., 1991). Percentage ASA (accessible surface area) is the percent of the total interface ASA contributed by each residue. Mutation nomenclature: E117S means Glu-117 was changed to serine. FcRn mutations (column 3) were made in rat FcRn, and effects on binding to rat IgG1 or IgG2a were assayed (Vaughn et al., 1997). Fc mutants in column 9 were introduced into human IgG1 or mouse IgG1 Fc regions, and effects on binding to mouse FcRn were assayed (Kim et al., 1994a, 1994b; Popov et al., 1996; Ghetie and Ward, 1997; Medesan et al., 1997; 1998) or introduced into a human IgG4 Fc region and evaluated using rat FcRn (Raghavan et al., 1995). Fc mutants in column 10 were introduced into human IgG1, and effects on binding to human FcRn were assayed (Shields et al., 2001). Effects listed as "relative binding" are expressed as a ratio of the binding of each mutant compared to wild-type IgG1, thus values less than 1.0 indicate reduced binding of the mutant compared to wild-type Fc, and values greater than 1.0 indicate enhanced binding of the mutant compared to wild-type Fc.

1995; Corper et al., 1997; DeLano et al., 2000) are compared. While the tertiary structures of individual Fc domains are relatively unchanged in the FcRn-bound hdFc, quaternary structural changes result in altered interdomain relationships. In unliganded Fc and Fc complexes containing two ligands, the Fc chains are related by an
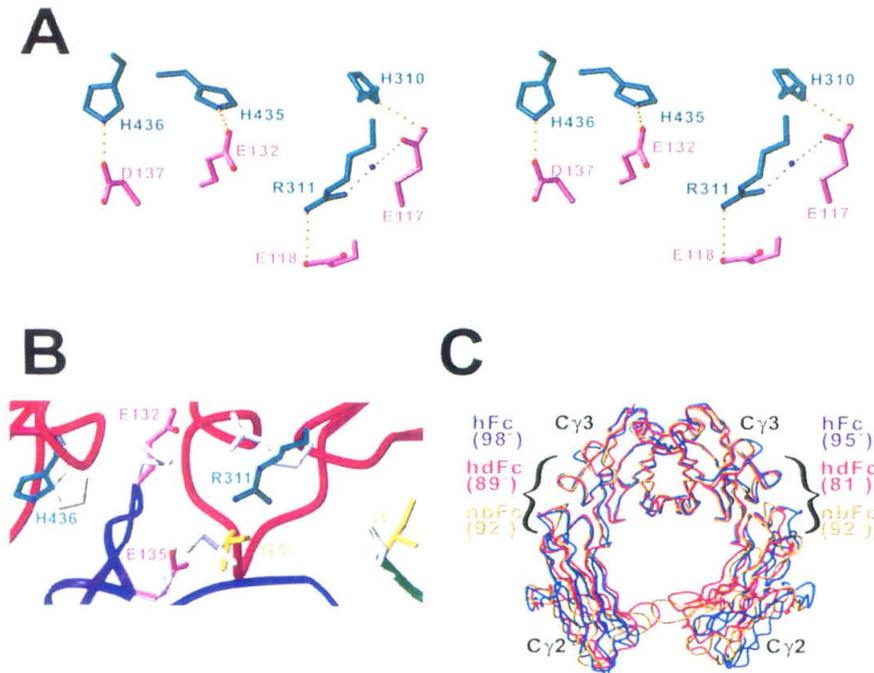
Figure 2. FcRn/hdFc Interface and Fc Interdomain Angles

(A) Salt bridge and hydrogen bond network at the interface between FcRn and Fc. Salt bridges are indicated by dotted yellow lines and were defined as interactions between oppositely charged atoms separated by less than 3.5 Å. Water-mediated hydrogen bonds are indicated in gray. Positively charged residues from Fc are turquoise; negatively charged residues from Fc are pink. The oxygen atom of an ordered water molecule is shown as a blue sphere.

(B) Comparison of residue conformations in the free and bound states of FcRn and Fc: FcRn from complex structure (turquoise), free FcRn (gray), Fc from complex structure (pink), free Fc (gray) (from the structures of human Fc [Deisenhofer, 1981] and nbFc), FcRn backbone (dark blue tubes), and Fc backbone (red tubes). Only interface residues that differ between the free and bound states are shown. In addition to changes at the interface region, loop regions of FcRn that are distant from the Fc binding site differ in backbone conformation when comparing the free and hdFc-bound FcRn structures. These changes, which do not appear to be induced by Fc binding, involve the loops between α2 domain strands 1 and 2 (residues 101–105, which includes an ordered N-linked carbohydrate not visible in previous structures [Burmeister et al., 1994a; Vaughn and Bjorkman, 1998]), α3 strands 1 and 2 (residues 190–195), and β2m strands 1 and 2 (residues 16–19) and 3 and 4 (residues 41–45). The latter three loops are involved in formation of the FcRn dimer observed in three crystal forms of free FcRn (Burmeister et al., 1994a; Vaughn and Bjorkman, 1998) but not in the FcRn/hdFc crystals.

(C) Comparison of Fc domain positions in hdFc, nbFc, and human Fc. The structures were aligned by superimposing a $C_\gamma 3$ domain on the proximal $C_\gamma 3$ domain of hdFc.

approximate dyad axis of symmetry, thus the $C_\gamma 2$–$C_\gamma 3$ interdomain angles (defined as described in Su et al., 1998) are identical or nearly the same for both chains (Table 2). However, in the FcRn/hdFc complex and other Fc complexes with only one ligand, the interdomain angles of the two chains differ, creating asymmetric Fc homodimers (Table 2). The range of interdomain angles in the free and ligand-bound Fc structures (85° to 106°) and the degree of interdomain angle asymmetry in singly liganded Fc structures (8°–13° different) demonstrate the intrinsic flexibility and the potential for asymmetry of Fc regions, which could have functional implications. For example, Fc asymmetry contributes to the 1:1 receptor/Fc stoichiometry observed in ligand binding by FcγRIII (Sondermann et al., 2000) and FcεRI (Garman et al., 2000). FcRn can also form 1:1 complexes with Fc under nonequilibrium conditions (Popov et al., 1996; Sánchez et al., 1999). The binding model for the FcRn/Fc interac-

tion assumes the following sequential reactions: FcRn + Fc ⇔ FcRn:Fc and FcRn:Fc + FcRn ⇔ FcRn:Fc:FcRn, and analyses of biosensor (Martin and Bjorkman, 1999) and equilibrium analytical ultracentrifugation data (Schuck et al., 1999) suggest that FcRn binding to the first binding site on an Fc homodimer lowers the affinity for binding a second FcRn. This could be explained on a structural level by Fc asymmetry such as that observed in the FcRn/hdFc structure. We cannot exclude the possibility that the observed hdFc asymmetry arises from crystal packing interactions or mutations introduced in the nbFc chain. We note, however, that the chains in the nbFc homodimer structure have a different $C_\gamma 2$–$C_\gamma 3$ interdomain angle (92°) than the angle in the nbFc chain of the hdFc structure (81°) (Table 2). Thus, the introduced mutations cannot be solely responsible for alteration of the nbFc interdomain angle in FcRn-bound hdFc structure.

Mechanism of pH-Dependent Binding

With the addition of the FcRn/hdFc structure to previous studies, we can now identify the residues involved in the sharp pH dependence of the FcRn/Fc interaction. In theory, pH-dependent binding and release could result from conformational changes in FcRn, Fc, or both. However, the structures of FcRn at pH 6.5 and 8 do not differ significantly (Vaughn and Bjorkman, 1998), and the pH 6.4 structure of nbFc reveals no major changes (other than those induced by mutations) when compared with structures of human Fc solved at other pH values: pH 6.5–7.0 (human Fc) (Deisenhofer, 1981), pH 4.1 (protein A/human Fc) (Deisenhofer, 1981), pH 7.8 (protein G/human Fc) (Sauer-Eriksson et al., 1995), pH 6.0 (human Fc/peptide) (DeLano et al., 2000), pH 5.6 (FcγRIII/human Fc) (Sondermann et al., 2000), pH 7.0 (rheumatoid factor/human Fc) (Corper et al., 1997), and pH 8.0 (intact mouse IgG2a) (Harris et al., 1997). The pH dependence of the FcRn/Fc interaction must therefore be attributed to chemical, rather than conformational, changes at the interface. Titration of histidines, which have a pK$_a$ in the range of the FcRn/Fc affinity transition, has been proposed to account for the pH dependence (Simister and Mostov, 1989; Ghetie and Ward, 2000). There are three pairs of titrating salt bridges at the FcRn/hdFc interface, each involving an Fc histidine and an acidic FcRn residue: Fc His-310/FcRn Glu-117, Fc His-435/FcRn Glu-132, and Fc His-436/FcRn Asp-137 (Figure 2A). The mechanism of the pH-dependent FcRn/Fc affinity transition appears straightforward: FcRn binds to Fc with high affinity at pH $\leq$ 6.5 when Fc histidines 310, 435, and 436 are positively charged and releases Fc upon deprotonation at pH values $\geq$7.0.

Although the FcRn/hdFc interface does not contain significantly more charged surface area than typical protein–protein interfaces (56% nonpolar, 29% polar, and 15% charged) (Lo Conte et al., 1999), the charged interface residues are involved in more salt bridges than seen in other Fc/protein complexes (Table 3). Formation of salt bridges between surface-exposed residues is not generally energetically favorable due to the loss of conformational entropy (Goldman, 1995). However, His-310 in unbound Fc and two of the three salt bridges at the FcRn/hdFc interface are buried (Fc His-310/FcRn Glu-117 and Fc His-435/FcRn Glu-132), and with the exception of Glu-132, the residues in these salt bridges assume the same conformations in the bound and free forms of the proteins. In addition, the residues in these salt bridges form part of a network of ionic interactions similar to those found on the surface of thermostable proteins (Goldman, 1995), such that Fc Arg-311 interacts with FcRn Glu-118 and an ordered water, which in turn contacts FcRn Glu-117 (Figure 2A). Salt bridge networks are energetically favorable because the cost of restricting the conformation of each additional residue is halved, while the coulombic benefit is essentially the same (Goldman, 1995). Since Fc Arg-311 is not conserved, the network of salt bridges observed in the rat FcRn/hdFc structure is not found in all FcRn/Fc complexes and must not be required for pH-dependent binding. In addition, the solvent-exposed salt bridge at the rat FcRn/Fc interface (Fc His-436/FcRn Asp-137) involves residues that vary in FcRn and Fc sequences (Table 3). These observations suggest that formation of

the two buried titratable salt bridges (Fc His-310/FcRn Glu-117 and Fc His-435/FcRn Glu-132) is sufficient to confer pH dependence to the FcRn/Fc interaction. Although other proteins contact Fc histidines at the C$_\gamma$2–C$_\gamma$3 interface, they do not show sharp pH-dependent Fc binding near neutral pH because they make fewer or no ionic interactions with Fc. For example, protein A and protein G do not form any salt bridges with Fc histidines (Deisenhofer, 1981; Sauer-Eriksson et al., 1995), and rheumatoid factor forms one relatively solvent-exposed salt bridge involving Fc His-433 (Corper et al., 1997).

Implications for Design of Fc-Containing Proteins with Increased Serum Half-Lives

Mutations that decrease the affinity between FcRn and IgG result in reduced serum half-lives in vivo (Kim et al., 1994a; Popov et al., 1996; Medesan et al., 1997), thus it has been suggested that mutant Fc regions with increased affinity for FcRn should exhibit increased serum persistence (Ghetie et al., 1997). Two studies have reported Fc mutants with increased affinity for FcRn. In the first, positions 252, 254, and 256 in a mouse IgG1 Fc fragment were randomly mutagenized, and one mutant (T252L, T254S, and T256F) showed an approximately 3.5-fold higher affinity for mouse FcRn and a longer β phase half-life when injected into mice (Ghetie et al., 1997). Fc residues 252 and 254 contact FcRn in the rat FcRn/hdFc structure (Table 3) and are likely to make similar contacts in a mouse FcRn/mouse Fc complex, rationalizing their effects upon affinity for FcRn. In the second study, alanine scanning mutagenesis was used to alter all solvent-exposed residues in a human IgG1 Fc, and binding to human FcRn was assessed (Shields et al., 2001). Mutants that showed higher or lower relative binding to FcRn than wild-type Fc are listed in Table 4, and their locations are depicted on the human IgG1 Fc structure (Deisenhofer, 1981) in Figure 3. To predict which residues on human Fc contact human FcRn, we constructed a model of a human FcRn/Fc complex by overlaying the structures of human FcRn (West and Bjorkman, 2000) and human Fc (Deisenhofer, 1981) on the rat FcRn/hdFc structure. The model must be regarded as approximate since some interface residues are not conserved between human and rat versions of FcRn and Fc (Table 3). Using the human FcRn/Fc complex model, we find that the majority of Fc mutations that result in diminished FcRn binding are in positions predicted to contact FcRn directly, whereas most of the mutations that enhance FcRn binding affect residues near but not actually inside of the region predicted to contact FcRn directly (Figure 3). Enhanced binding to FcRn by mutation of residues outside of the FcRn binding site suggests that propagation of small conformational changes can result in changes to the binding site that increase affinity, which may translate into increased serum persistence. Combination of mutants that individually have only a slight effect on FcRn affinity can result in 8- to 12-fold increases in relative FcRn binding (Table 4), suggesting that a synergistic effect on binding affinity can be achieved by this strategy.

From these results and an analysis of the FcRn/hdFc structure, we suggest a general strategy for identifica-

Table 4. Effect of Mutations in Human IgG1 Fc on Binding to Human FcRn

Mutation (Reduced Binding)	Relative Binding	Predicted to Contact FcRn?	Mutation (Enhanced Binding)	Relative Binding	Prediced to Contact FcRn?
I253A	<0.10	Yes	P238A	1.49	No
S254A	<0.10	Yes	T256A	1.91	No
R255A	0.59	No	E272A	1.34	No
K288A	0.38	Yes	V305A	1.46	No
L309A	0.63	Yes	T307A	1.81	Yes
S415A	0.44	No	Q311A	1.62	Yes
H433A	0.41	Yes	D312A	1.50	No
H435A	<0.10	Yes	K317A	1.44	No
Y436A	<0.10	Yes	D376A	1.45	No
			A378Q	1.32	No
			E380A	2.19	No
			E382A	1.51	No
			S424A	1.41	No
			N434A	3.46	Yes
			E380A + N434A	8.0	No/Yes
			T307A + E380A + N434A	11.8	Yes/No/Yes
			K288A + N434A	2.9	Yes/Yes

Fc mutants were introduced into human IgG1, and effects on binding to human FcRn were assayed (Shields et al., 2001). Effects listed as "relative binding" are expressed as a ratio of the binding of each mutant compared to wild-type IgG1, thus values less than 1.0 indicate reduced binding of the mutant compared to wild-type Fc, and values greater than 1.0 indicate enhanced binding of the mutant compared to wild-type Fc. Only substitutions resulting in relative binding values less than 0.70 or greater than 1.30 are listed. Wild-type residues at altered positions were evaluated for likelihood of contacting FcRn using a human FcRn/human Fc complex model generated from the rat FcRn/hdFc structure. To account for possible inaccuracies in the human FcRn/Fc model, the cutoff for contact residues was extended from 4 Å to 5 Å. Thus, Fc residues are predicted to contact FcRn if they contain an atom within 5 Å of FcRn.

tion of Fc mutants with increased affinity for FcRn. First, critical "functional epitope" residues (defined as $\Delta\Delta G >$ 2 kcal/mol for substitution of a single amino acid to

Figure 3. Positions that Affect Affinity for Human FcRn Highlighted on the Structure of Human Fc

A single polypeptide chain from the structure of human Fc (coordinates obtained from Mark Ultsch, Genentech) is shown with side chains highlighting positions where substitutions result in reduced (red side chain) or enhanced (green side chain) affinity for human FcRn, based upon mutagenesis studies by Shields et al. (2001) (Table 4). Residues within the predicted interface with human FcRn (within 5 Å of an FcRn atom using a human FcRn/human Fc model generated from the rat FcRn/hdFc structure) are indicated by thick side chains and labels. Residues predicted to be outside of the interface are indicated by thin side chains and smaller labels.

alanine, which corresponds to affinity reductions of >30-fold) (Cunningham and Wells, 1993) should remain unchanged. By this criterion, Ile-253, which contacts FcRn Trp 133, and His-310, which salt bridges with FcRn Glu-117 (Table 3), should not be altered. In addition, since FcRn-mediated rescue of IgG from catabolism requires pH-dependent binding (Ghetie and Ward, 2000), the sharp pH dependence of the FcRn/IgG interaction must be maintained in any mutants. This requires preserving His-435 as well as His-310. Although Fc His-436 in rat Fc also participates in a titrating salt bridge, the Fc His-436/FcRn Asp-137 salt bridge seen in our structure would be absent in a human FcRn/Fc complex, being replaced instead by a hydrophobic interaction involving an Fc tyrosine or phenylalanine interacting with FcRn Leu-137 (Table 3). Fc position 436 is therefore a candidate for random mutagenesis to find human Fc regions with higher affinity for FcRn. Other interface positions that could be substituted to produce human IgGs with longer half-lives include Fc residues 252, 254, 288, 307, 309, 311, 314, and 434, most of which have been modified by Ghetie et al. (1997) and/or Shields et al. (2001). Ideally, these residues would be targeted by a mutagenesis protocol allowing all possible non-glycine, non-proline substitutions at each position. A second category of candidates for mutagenesis includes residues that are near the FcRn binding site but do not make direct contact with FcRn. These residues include Fc positions 250, 251, 256, 257, 306, 308, 312, 431, 432, and 437. Alanine substitutions in many of these positions exhibited higher binding to human FcRn (Shields et al., 2001), thus an approach involving random substitutions at these positions may yield further increases in binding. A third category of potential mutants includes substitutions of residues buried between the $C_\gamma 2$ and $C_\gamma 3$ domains. Residues normally found at these positions allow

flexibility in the $C_\gamma2$–$C_\gamma3$ interdomain angle. However, optimal binding to FcRn is likely to require a particular $C_\gamma2$–$C_\gamma3$ interdomain angle from among a population of possible interdomain angles (e.g., see Table 2). Residues buried between the $C_\gamma2$ and $C_\gamma3$ domains include 247, 248, 251, 376, 378, 428, and 430. A mutagenesis strategy involving insertion of larger and more hydrophobic side chains at these positions could result in stabilization of the optimal $C_\gamma2$–$C_\gamma3$ interdomain angle for interaction with FcRn, thereby increasing the affinity for binding FcRn.

Discussion

The structure of the FcRn/Fc complex reported here allows a detailed analysis of FcRn binding to IgG, which increases our understanding of FcRn-mediated transport and protection of IgG and will benefit efforts to design therapeutic antibodies with longer serum half-lives. The hallmark of the FcRn/IgG interaction is the sharp pH dependence of the binding affinity that allows FcRn to associate with IgG in acidic intracellular vesicles and release IgG in the bloodstream. The FcRn/hdFc structure reveals a simple but elegant molecular mechanism for pH-dependent binding, which relies completely on titrating residues on the ligand rather than the receptor, such that Fc histidines interact favorably with negatively charged residues on FcRn at acidic but not basic pH. The finding that the pH-dependent binding interaction employed by FcRn is mediated entirely by titration of histidine side chains may be relevant to predicting the pH-dependent binding mechanisms of other receptor–ligand pairs that traffic through endosomes, such as transferrin receptor/HFE (Lebrón et al., 1998) and transferrin receptor/apo-transferrin (Richardson and Ponka, 1997) complexes. The pH-dependent affinity transition of the transferrin receptor/HFE interaction is in the opposite direction to that of FcRn/IgG (Lebrón et al., 1998), by contrast to that of transferrin receptor/apo-transferrin, which is the same direction as FcRn/IgG (Richardson and Ponka, 1997). There are two histidines at the interface between transferrin receptor and HFE (Bennett et al., 2000). Although neither histidine participates in a salt bridge, protonation at low pH could cause release of HFE from TfR. Alternatively, or perhaps concurrently, titration of histidines distant from the binding site, such as those at the transferrin receptor dimer interface, could result in a pH-dependent conformational change affecting binding of HFE and/or apo-transferrin (Bennett et al., 2000).

The FcRn/Fc structure also demonstrates the versatility of IgG recognition achieved by Fc receptors and other Fc binding proteins. Even among Fc binding proteins that bind to the $C_\gamma2$–$C_\gamma3$ interdomain interface of Fc, there are a variety of different folds, yet the proteins make many chemically similar contacts to the same residues on Fc (De Lano et al., 2000). FcRn differs from other Fc binding proteins in exhibiting homology to MHC molecules. The significance of the structural similarity between FcRn and MHC molecules is not clear, given that FcRn does not bind peptides or use the counterparts of the binding sites on class I MHC molecules or class I homologs for binding Fc. Instead, FcRn employs

a surface created by the juxtaposition of the $\beta2m$ and $\alpha2$ domains. Since the $\beta2m$/heavy chain interaction in FcRn is similar to that found in class I and class I homolog molecules (Burmeister et al., 1994a; Zeng et al., 1997; Lebrón et al., 1998; Strong, 2000), the common MHC heavy chain/$\beta2m$ interaction may predate the peptide and T cell–receptor binding functions of class I MHC proteins.

Experimental Procedures

Crystallization and Data Collection

Soluble rat FcRn (residues 1–269 of the heavy chain complexed with rat $\beta2m$) was purified from the supernatants of transfected CHO cells as described (Gastinel et al., 1992). Expression vectors encoding wtFc (rat IgG2a residues 223–447) and hdFc (IgG2a residues 223–447 with introduced mutations and a C-terminal factor Xa cleavage site and 6x-His tag) were cotransfected into CHO cells (Martin and Bjorkman, 1999). hdFc and nbFc were purified from supernatants of CHO cells secreting a mixture of wtFc, hdFc, and nbFc as described (Martin and Bjorkman, 1999). Briefly, CHO supernatants were passed over a Ni-NTA column, allowing separation of wtFc from 6x-His tagged species. hdFc and nbFc were eluted from the Ni-NTA column, then passed over an FcRn affinity column at pH 6.0. hdFc was eluted from the FcRn column at pH 8.0, and nbFc was recovered from the flowthrough.

Crystals of FcRn/hdFc complex (one 1:1 FcRn/hdFc complex per asymmetric unit) were grown in hanging drops containing 15% PEG 4000, 240 mM ammonium acetate, 120 mM sodium acetate (pH 4.5), and 1 mM $YtCl_2$. Crystal morphology was improved by serial rounds of streak-seeding. Improvements in crystal morphology accompanied improvements in crystal mosaicity (from 1.7° to 0.7°) and diffraction quality (from 8 Å to 2.8 Å). Crystals were transferred stepwise in three 5% increments to artificial mother liquor containing 15% 2-methyl-2,4-pentandiol, and 2.8 Å native data were collected at −170°C from flash-cooled crystals at SSRL beamline 9-2.

Crystals of nbFc (one nbFc homodimer per asymmetric unit) were grown in 0.2 M ammonium acetate, 0.1 M trisodium acetate dihydrate (pH 5.6), and 30% PEG 4000 and cryoprotected as described for the complex crystals. A 2.6 Å native data set was collected at −170°C using an R-AXIS IV mounted on a Rigaku RU-200 rotating anode generator.

Structure Determination and Refinement

Data were processed with the HKL package (Otwinowski and Minor, 1997). The structure of FcRn/hdFc was determined by molecular replacement using AmoRe (Navaza, 1994). The 2.2 Å structure of rat FcRn (Protein Data Bank ID code 3FRU) (Burmeister et al., 1994a; Vaughn and Bjorkman, 1998) and the 2.8 Å structure of human Fc (Protein Data Bank ID code 1FC1) (Deisenhofer, 1981) (nonconserved side chains truncated to alanine) were used as search models. FcRn was first located (correlation coefficient: 34.9%; R factor of 58.6%), then Fc was found in rotation and translation searches in which FcRn was fixed (correlation coefficient: 38.8%; R factor of 53.9%). Rigid body refinement (30–2.8 Å) of the two molecules as a complex resulted in an R_{cryst} of 43.1% (R_{free} = 42.2%). Solvent-flattened maps calculated to 3.0 Å were used for initial building using the program O (Jones and Kjeldgaard, 1997). Later stages of model building included use of higher-resolution (1.65 Å) human Fc structures provided by Mark Ultsch (Genetech) and the refined nbFc structure. For the nbFc structure solution, the truncated human Fc model was located using AmoRe (Navaza, 1994) (correlation coefficient: 23.7%; R factor of 51.5%), and rigid body refinement (30–2.7 Å) resulted in an R_{cryst} of 46.6% (R_{free} = 46.2%).

Refinement of the FcRn/hdFc and nbFc structures was done with CNS (Brünger et al., 1998) using overall anisotropic B factors, bulk solvent corrections, and tight NCS restraints (300 kcal/mol · Å²) for the $C_\gamma3$ with separate NCS operators for the $C_\gamma3$ and $C_\gamma2$ domains. For the FcRn/hdFc complex, hdFc regions that differ between the two Fc chains (chemically different side chains and loops at the FcRn contact site or in crystal contacts) were not restrained. Refinement of the FcRn/hdFc was conducted by alternating simulated

annealing and grouped B factor refinement. The electron density for most of the domains in the structure is excellent. However, the half of the distal $C_\gamma2$ domain furthest from the distal $C_\gamma3$ appears to be totally disordered. This portion of the structure was modeled using the nbFc coordinates that were placed into a simulated annealing omit electron density map (Hodel et al., 1992) in which the entire distal $C_\gamma2$ domain was omitted. Placement of the distal $C_\gamma2$ was facilitated by density for the disulfide between Cys-261 and Cys-321 and ordered carbohydrate at Asn-297. To determine whether disorder of the distal $C_\gamma2$ domain resulted from cryopreservation, data were collected at room temperature from FcRn/hdFc crystals using an R-AXIS IV image plate detector mounted on a Rigaku RU-200 rotating anode generator (R_{merge} = 0.19; 30–5.0 Å). The structure was solved by molecular replacement as described above (R_{cryst} = 0.35). $2F_o$-F_c maps showed disorder of the same region of the distal $C_\gamma2$ domain that is disordered in maps derived from the 2.8 Å data set collected from cryopreserved crystals.

The FcRn/hdFc model consists of residues 5–269 of FcRn, 1–99 of β2m, 239–443 of the proximal Fc (wtFc), 239–443 of the distal Fc (nbFc), six water molecules, 14 N-acetyl glucosamine, 4 fucose, and 8 mannose residues. Residues 1–4 of FcRn, the hinge regions of both Fc chains (residues 223–238), and the factor Xa cleavage site and 6x-His tag at the C terminus of the nbFc chain are disordered and not included in the model. Ordered carbohydrate is observed at all four potential N-linked glycosylation sites on FcRn and at the N-linked glycosylation site in each $C_\gamma2$ domain. The unpaired cysteines at positions 48 and 226 of FcRn are blocked by groups that were modeled as cysteine molecules. Average B values for individual domains are: FcRn α1–α2, 57 Å^2; FcRn α3, 64 Å^2; β2m, 60 Å^2; proximal Fc $C_\gamma3$, 82 Å^2; proximal Fc $C_\gamma2$, 55 Å^2; distal Fc $C_\gamma3$, 72 Å^2; distal Fc $C_\gamma2$, 147 Å^2; proximal Fc carbohydrate, 120 Å^2; and distal Fc carbohydrate 198 Å^2.

The nbFc model consists of two chains of residues 239–443 of rat IgG2a Fc, each containing 4 N-acetyl glucosamine, 1 fucose, and 3 mannose residues. Residues 223–238 (hinge region) and the C-terminal factor Xa cleavage site and 6x-His tag are disordered in both chains. Average B values per domain are: $C_\gamma2$ domains, 54 Å^2, 54 Å^2; $C_\gamma3$ domains, 23 Å^2, 24 Å^2; and carbohydrate, 91 Å^2, 100 Å^2.

Analyses of Structures

Buried surface areas were calculated from our structures and structures in the Protein Data Bank using a 1.4 Å probe radius with the programs GRASP (Nicholls et al., 1991) and Calc-surf (Gerstein, 1992) and verified using CNS (Brünger et al., 1998) and the Protein–Protein Interaction Server (Jones and Thornton, 1996). Interdomain angles were derived by determining the angle between the long axes of adjacent domains, approximated by ellipsoids calculated from the coordinates using the program Dom_angle (Su et al., 1998). Figures were made using Molscript (Kraulis, 1991) and rendered with Raster 3-D (Merritt and Murphy, 1994).

Coordinate superpositions were done using LSQMAN (Jones and Kjeldgaard, 1997). Rms deviations for domain-by-domain comparisons of FcRn alone and FcRn bound to hdFc: 0.6 Å (α1), 0.7 Å (α2), 0.5 Å (α3), and 0.5 Å (β2m). Rms deviations comparing hdFc domains with domains in six other structures (nbFc homodimer, free human Fc [Deisenhofer, 1981], human Fc/protein A [Deisenhofer, 1981], human Fc/protein G [Sauer-Eriksson et al., 1995], human Fc/rheumatoid factor [Corper et al., 1997], and human Fc/FcγRIII [Sondermann et al., 2000]): proximal $C_\gamma2$, 1.2 Å, 1.8 Å, 1.7 Å, 1.8 Å, 2.3 Å, 2.0 Å; proximal $C_\gamma3$, 0.4 Å, 0.7 Å, 0.7 Å, 0.7 Å, 0.6 Å, 0.8 Å; distal $C_\gamma2$, 1.6 Å, 2.3 Å, 2.3 Å, 2.3 Å, 2.8 Å, 2.6 Å; and distal $C_\gamma3$, 0.5 Å, 0.7 Å, 0.7 Å, 0.7 Å, 0.6 Å, 0.8Å.

Acknowledgments

References

Ahouse, J.J., Hagerman, C.L., Mittal, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Simister, N.E. (1993). Mouse MHC class-I like Fc receptor encoded outside the MHC. J. Immunol. 151, 6076–6088.

Bennett, M.J., Lebrón, J.A., and Bjorkman, P.J. (2000). Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor. Nature 403, 46–53.

Brünger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., et al. (1998). Crystallography and NMR system: a new software system for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54, 905–921.

Burmeister, W.P., Gastinel, L.N., Simister, N.E., Blum, M.L., and Bjorkman, P.J. (1994a). Crystal structure at 2.2 Å resolution of the MHC-related neonatal Fc receptor. Nature 372, 336–343.

Burmeister, W.P., Huber, A.H., and Bjorkman, P.J. (1994b). Crystal structure of the complex of rat neonatal Fc receptor with Fc. Nature 372, 379–383.

Chapman, T.L., You, I., Joseph, I.M., Bjorkman, P.J., Morrison, S.L., and Raghavan, M. (1999). Characterization of the interaction between the Herpes Simplex Virus type I Fc receptor and the immunoglobulin G. J. Biol. Chem. 274, 6911–6919.

Corper, A.L., Sohi, M.K., Bonagura, V.R., Steinitz, M., Jefferis, R., Feinstein, A., Beale, D., Taussig, M.J., and Sutton, B.J. (1997). Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction. Nat. Struct. Biol. 4, 374–381.

Cunningham, B.C., and Wells, J.A. (1993). Comparison of a structural and a functional epitope. J. Mol. Biol. 234, 554–563.

Deisenhofer, J. (1981). Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-Å resolution. Biochemistry 20, 2361–2370.

DeLano, W.L., Ultsch, M.H., de Vos, A.M., and Wells, J.A. (2000). Convergent solutions to binding at a protein-protein interface. Science 287, 1279–1283.

Garman, S.C., Wurzburg, B.A., Tarchevskaya, S.S., Kinet, J.P., and Jardetzky, T.S. (2000). Structure of the Fc fragment of human IgE bound to its high-affinity receptor Fc(epsilon)RI alpha. Nature 406, 259–266.

Gastinel, L.N., Simister, N.E., and Bjorkman, P.J. (1992). Expression and crystallization of a soluble and functional form of an Fc receptor related to class I histocompatibility molecules. Proc. Natl. Acad. Sci. USA 89, 638–642.

Gerstein, M. (1992). A resolution-sensitive procedure for comparing protein surfaces and its application to the comparison of antigen combining sites. Acta Crystallogr. A48, 271–276.

Ghetie, V., and Ward, E.S. (1997). FcRn - The MHC class I-related receptor that is more than an IgG transporter. Immunol. Today 18, 592–598.

Ghetie, V., and Ward, E.S. (2000). Multiple roles for the major histocompatibility complex class I-related FcRn. Annu. Rev. Immunol. 18, 739–766.

Ghetie, V., Popov, S., Borvak, J., Radu, C., Matesoi, D., Medesan, C., Ober, R.J., and Ward, E.S. (1997). Increasing the serum persistence of an IgG fragment by random mutagenesis. Nat. Biotechnol. 15, 637–640. .

Goldman, A. (1995). How to make my blood boil. Structure 3, 1277–1279.

Harris, L.J., Larson, S.B., Hasel, K.W., and McPherson, A. (1997). Refined structure of an intact IgG2a monoclonal antibody. Biochemistry 36, 1581–1597.

Hodel, A., Kim, S.-H., and Brünger, A.T. (1992). Model bias in macromolecular crystal structures. Acta Crystallogr. *A48*, 851–858.

Janin, J. (1997). Specific versus non-specific contacts in protein crystals. Nat. Struct. Biol. *4*, 973–974.

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

Jones, T.A., and Kjeldgaard, M. (1997). Electron density map interpretation. Methods Enzymol. *277*, 173–208.

Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S., and Foeller, C. (1991). Sequences of Proteins of Immunological Interest (Bethesda, MD: U.S. Department of Health and Human Services).

Kacskovics, I., Wu, Z., Simister, N.E., Frenyó, L.V., and Hammarström, L. (2000). Cloning and characterization of the bovine MHC class I-like Fc receptor. J. Immunol. *164*, 1889–1897.

Kim, J.-K., Tsen, M.-F., Ghetie, V., and Ward, E.S. (1994a). Localization of the site of the murine IgG1 molecule that is involved in binding to the murine intestinal Fc receptor. Eur. J. Immunol. *24*, 2429–2434.

Kim, J.K., Tsen, M.F., Ghetie, V., and Ward, E.S. (1994b). Identifying amino acid residues that influence plasma clearance of mouse IgG1 fragments by site directed mutagenesis. Eur. J. Immunol. *24*, 542–548.

Kim, J.-K., Firan, M., Radu, C.G., Kim, C.-H., Ghetie, V., and Ward, E.S. (1999). Mapping the site on human IgG for binding of the MHC class I-related receptor, FcRn. Eur. J. Immunol. *29*, 2819–2825.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J. Mol. Biol. *234*, 946–950.

Lebrón, J.A., Bennett, M.J., Vaughn, D.E., Chirino, A.J., Snow, P.M., Mintier, G.A., Feder, J.N., and Bjorkman, P.J. (1998). Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. Cell *93*, 111–123.

Lo Conte, L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. J. Mol. Biol. *285*, 2177–2198.

Martin, W.L., and Bjorkman, P.J. (1999). Characterization of the 2:1 complex between the class I MHC-related Fc receptor and its Fc ligand in solution. Biochemistry *38*, 12639–12647.

Medesan, C., Matesoi, D., Radu, C., Ghetie, V., and Ward, E.S. (1997). Delineation of the amino acid residues involved in transcytosis and catabolism of mouse IgG. J. Immunol. *158*, 2211–2217.

Medesan, C., Cianga, P., Mummert, M., Stanescu, D., Ghetie, V., and Ward, E.S. (1998). Comparative studies of rat IgG to further delineate the Fc/FcRn interaction site. Eur. J. Immunol. *28*, 2092–2100.

Merritt, E.A., and Murphy, M.E.P. (1994). Raster3D Version 2.0, a program for photorealistic molecular graphics. Acta Crystallogr. D. *50*, 869–873.

Navaza, J. (1994). AMORE— an automated package for molecular replacement. Acta Crystallogr. *A50*, 157–163.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Popov, S., Hubbard, J.G., Kim, J.-K., Ober, B., Ghetie, V., and Ward, E.S. (1996). The stoichiometry and affinity of interaction of murine Fc fragments with the MHC class I-related receptor, FcRn. Mol. Immunol. *33*, 521–530.

Raghavan, M., Bonagura, V.R., Morrison, S.L., and Bjorkman, P.J. (1995). Analysis of the pH dependence of the neonatal Fc receptor/Immunoglobulin G interaction using antibody and receptor variants. Biochemistry *34*, 14649–14657.

Richardson, D.R., and Ponka, P. (1997). The molecular mechanisms of the metabolism and transport of iron in normal and neoplastic cells. Biochim. Biophys. Acta *1331*, 1–40.

Sánchez, L.M., Penny, D.M., and Bjorkman, P.J. (1999). Stoichiometry of the interaction between the MHC-related Fc receptor and its Fc ligand. Biochemistry *38*, 9471–9476.

Sauer-Eriksson, A.E., Kleywegt, G.J., Uhlen, M., and Jones, T.A. (1995). Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. Structure *3*, 265–278.

Schuck, P., Radu, C.G., and Ward, E.S. (1999). Sedimentation equilibrium analysis of recombinant mouse FcRn with murine IgG1. Mol. Immunol. *36*, 1117–1125.

Shields, R.L., Namenuk, A.K., Hong, K., Meng, Y.G., Rae, J., Briggs, J., Xie, D., Lai, J., Stadlen, A., Li, B., et al. (2001). High resolution mapping of the binding site on human IgG1 for Fc gamma RI, Fc gamma RII, Fc gamma RIII, and FcRn and design of IgG1 variants with improved binding to the Fc gamma R. J. Biol. Chem. *276*, 6591–6604.

Simister, N.E., and Mostov, K.E. (1989). An Fc receptor structurally related to MHC class I antigens. Nature *337*, 184–187.

Sondermann, P., Huber, R., Oosthuizen, V., and Jacob, U. (2000). The 3.2 Å crystal structure of the human IgG1 Fc fragment-FcγRIII complex. Nature *406*, 267–273.

Story, C.M., Mikulska, J.E., and Simister, N.E. (1994). MHC class I-like Fc receptor cloned from human placenta. J. Exp. Med. *180*, 2377–2381.

Strong, R.K. (2000). Class (I) will come to order—*not*. Nat. Struct. Biol. *7*, 173–176.

Su, X.-D., Gastinel, L.N., Vaughn, D.E., Faye, I., Poon, P., and Bjorkman, P.J. (1998). Crystal structure of hemolin: a horseshoe shape with implications for homophilic adhesion. Science *281*, 991–995.

Vaughn, D.E., and Bjorkman, P.J. (1998). Structural basis of pH-dependent antibody binding by the neonatal Fc receptor. Structure *6*, 63–73.

Vaughn, D.E., Milburn, C.M., Penny, D.M., Martin, W.L., Johnson, J.L., and Bjorkman, P.J. (1997). Identification of critical IgG binding epitopes on the neonatal Fc receptor. J. Mol. Biol. *274*, 597–607.

West, A.P., and Bjorkman, P.J. (2000). Crystal structure and IgG binding properties of the human MHC-related Fc receptor. Biochemistry *39*, 9698–9708.

Zeng, Z.H., Castaño, A.R., Segelke, B., Stura, E.A., Peterson, P.A., and Wilson, I.A. (1997). The crystal structure of murine CD1: an MHC-like fold with a large hydrophobic binding groove. Science *277*, 339–345.

Protein Data Bank ID Codes

The FcRn/hdFc and nbFc coordinates have been deposited in the Protein Data Bank under ID codes 1I1A and 1I1C, respectively.

Chapter 4:

Examination of the ability of the neonatal Fc receptor

to form a receptor dimer

This chapter covers the various studies that were conducted to ascertain whether FcRn forms a dimer of heterodimers. These studies include quantitative analyses of the FcRn dimer interface present in FcRn crystals and sedimentation equilibrium and dynamic light scattering studies of the soluble FcRn ectodomain performed by our collaborator, Dr. Ghirlando Rudolfo, at the NIH. We also performed experiments that attempted to address the behavior of FcRn in the membrane. The first set of experiments involving membrane bound FcRn, ligand-induced cross-phosphorylation assays, was performed with Dr. T. S. Ramalingham, of our lab. These employed a chimera of the FcRn ectodomain and a dimerization reporter group consisting of cytoplasmic tyrosine kinase domain of TrkA constructed by Dr. Yang Liu. The second set of experiments, measurements of fluorescent resonance energy transfer, was performed with Dr. Mary Dickinson of the Biological Imaging Center. These employed a set of chimeras of the full-length FcRn protein with the cyan or yellow forms of green fluorescent protein fused to the carboxy-terminus of the FcRn cytoplasmic domain.

Introduction

Previous studies of the neonatal Fc receptor (FcRn) have accumulated evidence that FcRn forms a dimer under some conditions. First, electron irradiation studies conducted on the FcRn isolated from the gut of neonatal rats suggested that the functional molecular weight is 110 kDa, which corresponds to two FcRn molecules (Simister and Rees, 1985). Second, in three crystal forms FcRn forms the same dimer (Burmeister et al., 1994a). The same FcRn dimer is also observed in the low-resolution cocrystal structure (~6.5 Å) of a rat FcRn ectodomain/Fc complex, in which FcRn dimers were bridged by Fc molecules (Burmeister et al., 1994b; See Appendix A for a description of the Fcrn dimer). Each of the two potential FcRn binding sites on Fc interacts with one of the FcRn molecules in the receptor dimer, resulting in a long oligomeric ribbon where there are two receptors for every Fc dimer (Burmeister et al., 1994b). Surface plasmon resonance assays of FcRn molecules with site-directed mutations of positions at the crystallographic FcRn dimer interface (FcRn Gly191 and β2m Glu89) revealed that these mutant forms had lower affinities for Fc and IgG. These data suggested that receptor dimerization was necessary for high affinity binding of IgG, consistent with the receptor dimers being required for biological function (Vaughn et al., 1997; Appendix 1).

Surface plasmon resonance, immuno-precipitation, isothermal titration calorimetery, and equilibrium gel filtration studies with the soluble FcRn ectodomain and the Fc portion of IgG showed that the stoichiometry is two receptors per ligand (2:1) in solution (Vaughn et al., 1997; Huber et al., 1993; Sanchez and Bjorkman, 1999). This stoichiometry is also seen with the receptor anchored to solid support and in crystals (Vaughn et al., 1997; Huber et al., 1993; Sanchez and Bjorkman, 1999). The oligomeric ribbon presents two distinct complexes that have a 2:1 FcRn-Fc stoichiometry: Fc bound to the FcRn dimer observed in the FcRn crystals and the bridging of two separate FcRn molecules by Fc (See chapter 2 for a depiction of the oligomeric ribbon). To determine

which of the possible 2:1 FcRn Fc complexes forms in these conditions, a heterodimeric Fc (hdFc) molecule with one FcRn binding site was constructed. This molecule could bind a dimer of FcRn molecules but it could not bridge FcRn molecules (See Chapter 3). In surface plasmon resonance, immuno-precipitation, and equilibrium gel filtration studies this molecule bound to FcRn with 1:1 FcRn:hdFc stoichiometry (Martin and Bjorkman, 1999; Chapter 3). These data indicated that the 2:1 FcRn:Fc complex formed in solution and on a solid support was two separate FcRn molecules bridged by an Fc.

In addition to the determination that an Fc-induced dimer of FcRn molecules was not observed in solution or with FcRn coupled to a solid support, other studies of FcRn indicated the FcRn dimer only forms at high concentrations of FcRn. In the $C222_1$ crystal form of rat FcRn at both acidic and alkaline pH, there are three FcRn molecules in the asymmetric unit and only two of the three FcRn molecules are in an FcRn dimer (Vaughn and Bjorkman, 1998). Both the crystal structure of human FcRn alone (West and Bjorkman, 2000) and the crystal structure of the rat FcRn/hdFc complex (Martin et al., 2001; Chapter 4) do not contain FcRn dimers. The FcRn concentration in the crystals is millimolar. The presence of an FcRn dimer in some crystals and the absence in other crystals suggests that the dissociation constant for dimer formation is on the order of millimolar. For the FcRn dimer to occur in living systems, it is reasonable to infer either or both of the following: the effective molarity of FcRn in a membrane is higher than millimolar, the affinity of full-length FcRn for itself is higher than that of the ectodomains alone.

In order to address more definitively under what conditions FcRn forms an FcRn dimer, two different sets of investigations were pursued: the propensity of the FcRn ectodomain to form a dimer and the tendency of FcRn to form a dimer in a membrane. For the studies of the FcRn ectodomain, the FcRn-FcRn contacts in the FcRn crystals were analyzed to determine if they more resemble a nonspecific crystal contact or a

physiologically relevant dimer interface. In addition we colloborated with Ghirlando Rudolfo of the NIH to perform sedimentation equilibrium and dynamic light scattering experiments to detect dimer formation with the soluble ectodomain at high concentrations. To examine the behavior of membrane-bound forms of FcRn two sets of FcRn-reporter group fusion proteins were made. The first set of fusion proteins consisted of the FcRn ectodomain fused to the transmembrane and cytoplasmic tyrosine kinase domains of TrkA. These chimeric proteins were assayed for ligand-induced cross-phosphorylation as a read-out for dimerization. The second set of fusion proteins created was full-length FcRn fused to ECFP or EYFP. These chimeric proteins were constructed and assayed for fluoresence resonance energy transfer (FRET). The results from the FcRn ectodomain studies show only weak resemblance on the part of the crystallographic contacts to a physiologically relevant dimer interface and no detectable formation of dimer by the ectodomain alone in solution. The studies involving membrane bound forms of FcRn are more difficult to interpret. Some experiments have indicated that FcRn may form a ligand-induced dimer or may be a dimer constituitively. However, these results are not consistently reproduced and their biological significance is difficult to assess. Strategies are discussed for constructing a system in which FcRn dimer formation may be detected and its functional importance addressed.

Results

Analysis of the FcRn/FcRn crystal contact

The crystal structures of rat FcRn at acidic pH (Burmeister et al., 1994a) and at alkaline pH (Vaughn and Bjorkman, 1998) reveal a large area of contact between two of the three receptor heterodimers. As seen in Figure 1, each FcRn molecule buries 890 Å^2, in which the α3 domains of both receptors contact the α3 and β2m domains of the other receptor. The approximately two-fold symmetric interaction involves the loops between

the A and B, and E and F strands of the α3 domain as well as the G strand of β2m. In both the acidic and the alkaline pH FcRn structures this contact is enlarged considerably (from 1500 Å^2–1800 Å^2 to 2600 Å^2) by the ordered carbohydrate extending from Asn 225 of the FcRn α3 domain across the other FcRn molecule (Vaughn and Bjorkman, 1998).

It is difficult to distinguish between a biologically meaningful and a nonspecific contact between proteins in the crystal, however there are several ways to analyze the structures to indicate the likelihood of whether a given contact is nonspecific or represents a biologically relevant interaction. The first method involves an analysis of the amount of surface area buried by the contact (Janin, 1997). The second method involves examining how well the proteins fit together (Lawrence and Colman, 1993). The third method involves determining if the surfaces in contact are associated with sequence that is more conserved than the sequence associated with other areas of the proteins surface (Elcock and McCammon, 2001).

The method developed by Janin for judging whether crystal contacts represent a meaningful interaction is based on the observation that the larger the contact between proteins the more likely the contact is to be biologically significant. Janin compiled a database of 1320 pair-wise interfaces of 152 crystal forms of monomeric proteins with only one molecule per asymmetric unit (Janin and Rodier, 1995). A histogram of the frequency of interfaces of a given size is approximated by an exponential function. For contacts between proteins greater than 700 Å^2 in size, the probability of finding a nonspecific contact between monomeric proteins with the surface area in question or greater is equal to the area under the exponential decay from the surface area to infinity. The solution of this integral is the formula for the probability:

$$P(B) \sim 4.2 \exp(-B/260), \tag{1}$$

where B is the amount of surface area buried at the contact on both proteins. FcRn buries 1780 Å^2 total at the dimer interface (without considering the contribution of the ordered carbohydrate), by this method, the probability of a nonspecific contact between monomers being this size or larger is 0.4%. It must be noted that the FcRn crystal has three FcRn molecules per asymmetric unit and that the data set used in this method was constructed to include only crystal structures with one molecule per asymmetric unit. It should also be noted that monomeric molecules in crystals with point group symmetries, such as a crystallographic two-fold, often have contacts that are larger than those without (Janin and Rodier, 1995; Carugo and Argos, 1997). There are too few of these structures to generate a distribution that may be fit to a function with confidence. The FcRn contact in question is the interface about a non-crystallographic two-fold in the high resolution FcRn crystals, and this same contact occurs about a crystallographic two-fold in the FcRn/Fc complex crystals. For these reasons it is unclear how much predictive power this algorithm has for the contact in FcRn crystals.

The second method for analyzing the FcRn-FcRn crystal contact is to measure the degree of fit between the FcRn molecules. Though it has not been examined systematically, it is intuitively satisfying to imagine that specific biological contacts have a higher degree of fit than nonspecific contacts. There are several ways to assess the degree of fit between two molecules. We chose to use the shape complementarity statistic (Lawrence and Colman, 1993). The shape complementarity statistic samples points on the surface between two proteins and assesses the distance between the points and the angle between the vectors normal to the surfaces at these points. The shape complementarity statistic that ranges from 0 to 1, where 1 is perfect complementarity, is

the median value of these samples. The shape complementarity of the FcRn-FcRn contact is 0.53, significantly lower than the average shape complementarity seen in nonobligate protein complexes (Appendix B). However, this is only suggestive of the interface being nonspecific. There are specific interfaces with shape complementarity this low, Protein G and Fc is 0.56 (Chapter 3), and there are nonspecific contacts with very high shape complementarity. The crystal structure of Fe-only hydrogenase (Peters et al., 1998), which is a monomer in solution, has a crystal contact of large surface area (1575 Å^2) (Elcock and McCammon, 2001). The shape complementarity of this contact is found here to be 0.75. The shape complementarity of the contact between FcRn molecules indicates a low degree of fit; however, in the absence of a systematic review it is not clear how diagnostic this low degree of fit is to the biological significance.

Recently, Elcock and McCammon demonstrated a third method for examining the biological significance of a crystal contact that employs protein sequence information as well as the crystal structure (Elcock and McCammon, 2001). The tendency of the sequence to vary at a given position is the disorder or entropy of the sequence at this position (defined mathematically in the methods). This method maps the tendency for the sequence to vary at a given position in sequence to the position on the surface of the structure by weighting the sequence entropy of a given residue by the amount of solvent accessible surface area associated with that residue. The average surface entropy for the surface buried in the contact is divided by the average surface entropy for the rest of the surface. If this ratio is less than one, the contact is considered to be biologically significant. To measure the sequence entropy of FcRn we used the 5 known FcRn sequences (rat, mouse, opossum, bovine and human) and 28 sequences of β2

microglobulin. While the number of β2 microglobulin sequences is enough for this analysis (>20), the number of FcRn sequences is probably too low and may result in an overstatement of the sequence entropy (Elcock and McCammon, 2001). The surface sequence entropy for the FcRn-FcRn contact is slightly higher than the surface sequence entropy for the surface of FcRn generally. The 1.11 ratio is consistent with interpreting the FcRn dimer as a crystal contact and not as a biologically meaningful contact. However, the limiting number of FcRn sequences diminishes the significance of this estimate.

Analyzing the FcRn-FcRn crystal contact computationally generates results that are not straightforward to interpret. The absence of a clear indication that the FcRn-FcRn crystal contact is biologically meaningful left any demonstration of its existence to direct experiments. Ghirlando Rudolfo (NIH) conducted the first set of experiments using the soluble FcRn ectodomain. The second set was performed on a chimeric protein expressed in CHO cells in which the soluble FcRn ectodomain was fused to the transmembrane and cytoplasmic domains of TrkA, a tyrosine kinase. The final set of experiments employs FcRn-ECFP and FcRn-EYFP chimeras constructed with the entire FcRn heavy chain with the enhanced yellow (or cyan) fluorescent protein fused carboxy-terminus to the short (43 amino acid) cytoplasmic tail of FcRn.

Analytical Ultracentrifugation and dynamic light scattering of the soluble FcRn ectodomain

Ghirlando Rudolfo (NIH) conducted the following experiments on our behalf with soluble FcRn we supplied him. These results are included for the purpose of presenting a more complete discussion of the relevant unpublished data.

Sedimentation equilibrium experiments were carried out in order to determine the oligomeric state of FcRn in 0.15M NaCl and 0.063M sodium phosphate (pH = 6.0). At a loading concentration of 10.7 μM sedimentation equilibrium gradients were best modeled in terms of a single ideal solute (see Materials and Methods: Equation [5]). Identical values of $M(1 - v\rho)$ were obtained at the different rotor speeds, averaged at $12,900 \pm 140$ g mol^{-1}, showing that the sample is monodisperse (Figure 1).

The experimental molecular mass, M, was calculated using equation [7] as described in (Ghirlando et al. 1995). Based on the amino acid composition of the FcRn components, a protein molecular mass M_p of 41,930 g mol^{-1} is calculated. Similarly, based on the consensus data published by Perkins (Perkins, 1986), a protein partial specific volume v_p of 0.7328 mL g^{-1} is determined. Assuming that the carbohydrate has a partial specific volume v_c of 0.650 mL g^{-1} (Ghirlando et al., 1995; Durschlag, 1986), the experimentally determined molecular mass is calculated as $47,500 \pm 510$ g mol^{-1}, showing that the FcRn is monomeric. These data are consistent with a 12% (w/w) glycosylation and are in reasonable agreement with the molecular mass of 49,260 g mol^{-1} determined by mass spectrometry.
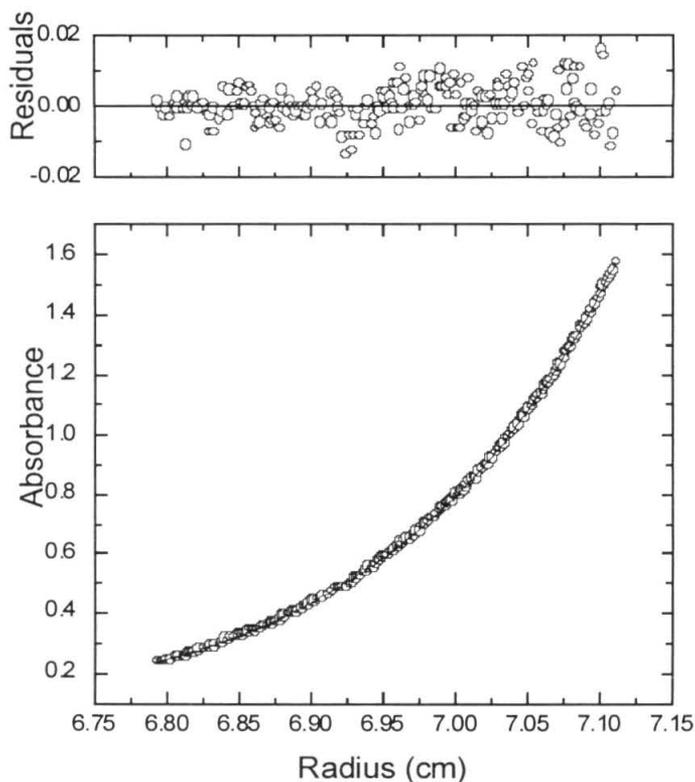
Figure 1: FcRn is a single, monodisperse monomer in 0.15M NaCl and 0.063M sodium phosphate (pH = 6.0). Sedimentation equilibrium profile at 280 nm for FcRn at a loading concentration of 10.7 μM, shown as a distribution of A_{280} at equilibrium. Data were collected at 4.0° C and 12,000 rpm. The results are analyzed for the best single component $M(1 - v\rho)$ fit, shown as a line through the experimental points. The corresponding distribution of the residuals is shown above the plot.

Table 1

Buoyant molecular mass as a function of the FcRn loading concentration. [a]

[FcRn] / mM	λ / nm	Rotor speeds / 1,000 rpm	$M(1 - v\rho)$ / g mol^{-1}
0.0107	280	10, 12, 14	12,900 ± 140
0.0214	298	10, 12, 14	12,870 ± 350
0.0429	302	10, 12, 14	12,880 ± 120
0.0857	304	10, 12, 14	12,350 ± 110
0.105	305	8, 10	12,330 ± 130
0.172	306	10, 12, 14	11,870 ± 360
0.210	305	8, 10	11,600 ± 40
0.420	310	8 only[b]	10,760 ± 50
1.00	365	6 only[b]	7,720 ± 80

(a) Experimental parameters for the determination of the buoyant molecular mass. The values of $M(1 - v\rho)$ shown, obtained from the best fit of equation [5] to the experimental data, represent an average of the values determined at the different rotor speeds.

(b) At these loading concentrations, spinning at high speeds leads to FcRn concentrations that scatter light. Data could not be collected at higher rotor speeds, values of $M(1 - v\rho)$ were determined at a single rotor speed.

In order to determine whether the FcRn self-associates into higher order oligomers, sedimentation equilibrium experiments were carried out at a series of higher loading concentrations (Table 1). Even though the experimental data were always consistent with the presence of a monodisperse species, it was noted that the values of $M(1 - v\rho)$ decrease monotonically with increasing concentration (Table 1, Figure 2). In fact, it was found that the dependence of $M(1 - v\rho)$ upon the FcRn concentration in g L^{-1} (w_{FcRn}) was best described by the following linear relationship (Figure 2):

$$M(1 - v\rho) = 12,900 - 110\ w_{FcRn} \qquad\qquad [A]$$

**FcRn buovant molecular mass as a
function of the FcRn concentration**
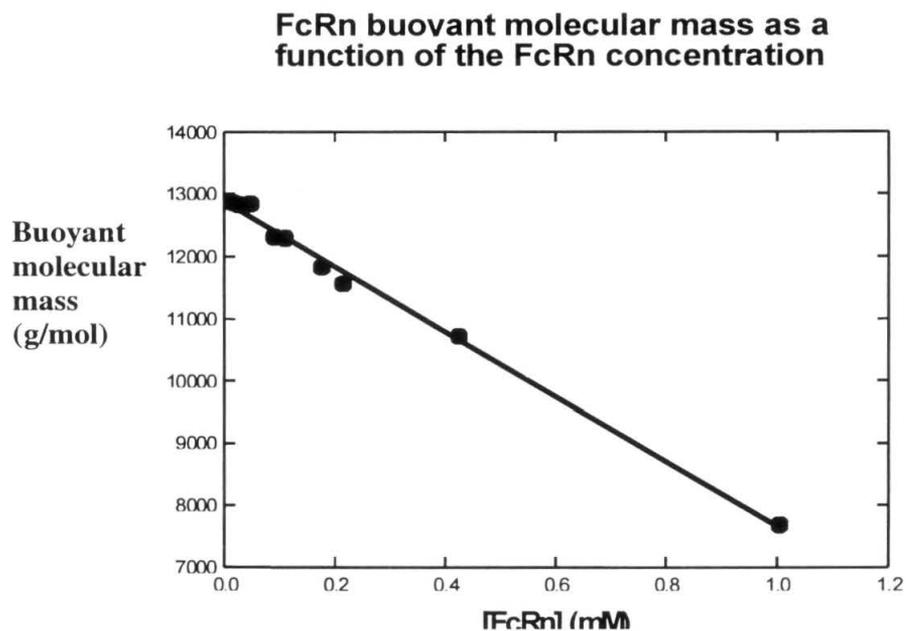
Figure 2: Buoyant molecular mass of FcRn in 0.15M NaCl and 0.063M sodium phosphate (pH =

6.0) as a function of the FcRn concentration. The line shows the best-fit linear analysis of the

data. Based on the experimental molecular mass of 47,500 g mol⁻¹, (in order to convert the

concentration scale from [FcRn] in mM to w_{FcRn} in g L⁻¹) the best-fit line is described by equation

[3] in the text.

A similar monotonic dependence of the buoyant molecular mass on the sample concentration has documented for monomeric BSA and monomeric fibrinogen (Rivas et al., 1999). In these cases, the decrease of the apparent molecular mass was interpreted in terms of volume exclusion (Rivas et al., 1999; Minton, 1998). Data for the FcRn can be interpreted in a similar manner: the values of the activity coefficient γ for the FcRn were calculated using equation [A] above and equation [6] described in Rivas et al. (Rivas et al., 1999) (Figure 4). Treatment of the FcRn solution as a suspension of hard spherical particles (Equation [10] in Rivas et al., 1999), shows that the variation of $\ln\gamma$ is best modeled in terms of a FcRn having a radius $r_{FcRn} = 2.8$ nm (Figure 3).

In order to verify the validity of this model, FcRn samples were studied by dynamic light scattering which allows for an independent measure of the effective particle radius in solution. As in the sedimentation equilibrium experiments, it was noted that the particle diffusion coefficient and radius depended on the sample concentration (Figure 4). This is due to a combination of volume exclusion and multiple scattering events. Extrapolation to a zero concentration using a simple quadratic equation leads to an extrapolated radius of 2.7 ± 0.1 nm, a value which is identical to that determined in Figure 4.

Therefore, the monotonic decrease of the molecular mass as a function of the FcRn concentration indicates that the FcRn does not form dimers or higher oligomers at the concentrations studied. The formation of such species would have been evidenced by a 'break' in the linear dependence of the experimental molecular mass, as noted for fibrinogen in the presence of calcium ions (Figure 2 of Rivas et al., 1999). This dependence of the molecular mass decrease is a thermodynamic consequence of

Figure 3: Activity coefficients of FcRn calculated from the experimental data (equation [5]) and equation [6] in reference (4). Integrations were carried out numerically in Mathematica 4.0 (Wolfram Research) and the particle number density ρ_0 in mL^{-1} was calculated as 10^{-6}[FcRn]N_a, where N_a is Avogadro's number and [FcRn] is the concentration in mM. The line shows the best-fit scaled particle theory based on hard particle FcRn spheres.

**FcRn radius as a function
of concentration**
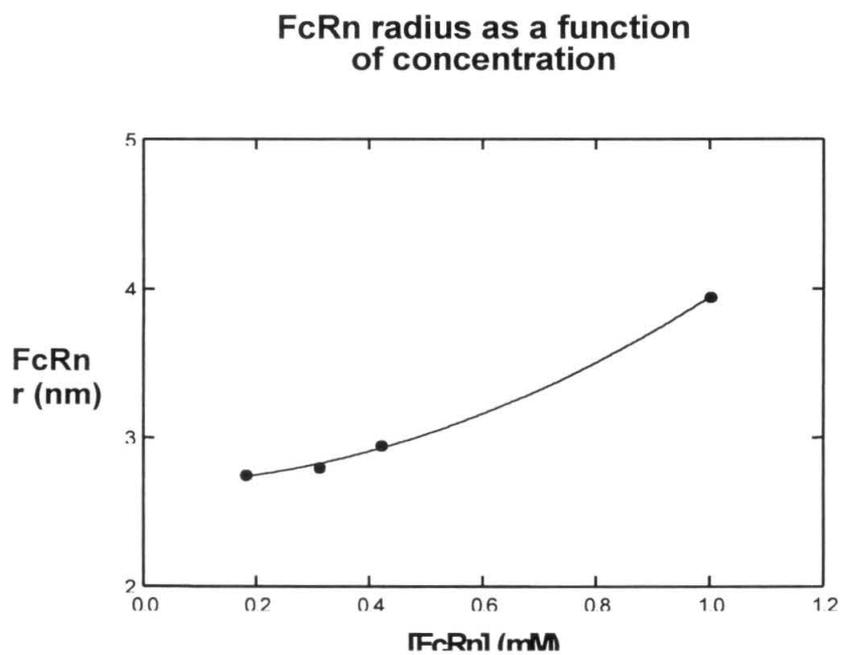
Figure 4: FcRn radius determined by dynamic light scattering as a function of the sample concentration. The line shows the quadratic best-fit.

molecular crowding. Analysis of the data in terms of a hard particle model is consistent with dynamic light scattering experiments, further supporting the conclusion that the FcRn does not self-associate under the conditions studied.

In summary, sedimentation equilibrium and dynamic light scattering experiments on the soluble ectodomain of FcRn both indicate that soluble FcRn is a monomer in solution at concentrations up to 1.00 millimolar.

Cross-phosphorylation assays of an FcRn-TrkA chimeras

One way to determine whether FcRn formed a dimer while in a membrane was to attach a FcRn to a group that would report formation of a dimer. Dr. Yang Liu, a former graduate student in our lab, constructed the first such FcRn-reporter protein by attaching the tyrosine kinase domain of the high affinity nerve growth factor receptor, TrkA, to the ectodomain of FcRn. The FcRn-TrkA chimera includes the ectodomain of rat FcRn and the transmembrane domain and cytoplasmic domain of rat TrkA. Wild-type TrkA undergoes ligand induced cross-phosphorylation in the presence of nerve growth factor (Farooqui et al., 1997). Fusing the C-terminal domains of TrkA to FcRn thus couples a dimer-reporting function to the Fc binding domain of FcRn. An expression vector containing the gene for FcRn-TrkA was transfected into Chinese hamster ovary (CHO-K1) cells, stable transfectants were isolated by cell sorting using antibodies recognizing FcRn and rat $\beta2$ microglobulin (Liu and Bjorkman, unpublished results). The transfected cells were shown to bind iodinated Fc (Figure 5).
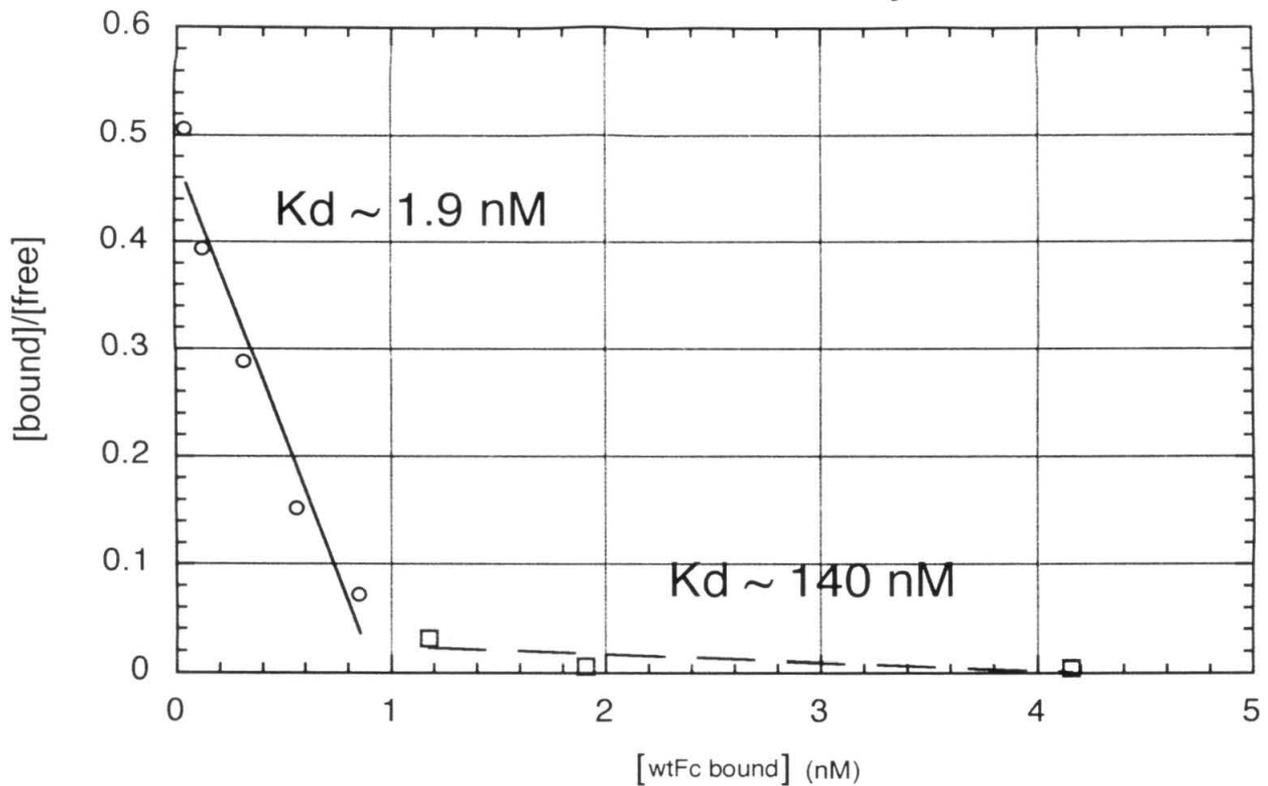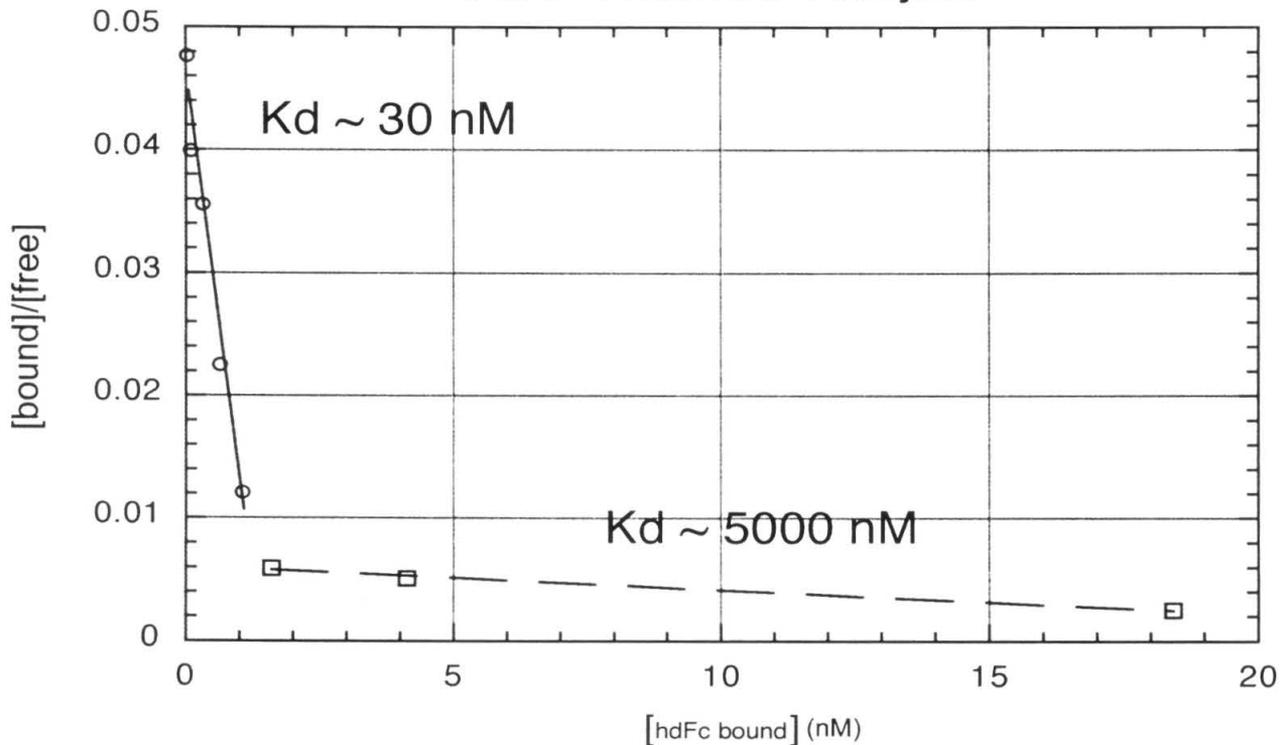
Figure 5: Binding of wtFc and hdFc to FcRn-TrkA expressing cells. Biphasic analysis adopted from Vaughn & Bjorkman (1997). Dissociation constants are consistent with values generated in SPR studies of FcRn and the ligandswhen analyzed with the biphasic analysis (Martin & Vaughn, unpublished results).

In collaboration with Dr. T. S. Ramalingham, the cells were assayed for the induction of cross-phosphorylation in the presence and absence of added ligand (Figure 6). The cells were serum-starved then serum and phosphate starved and then incubated with $H_3{}^{32}PO_4$. Afer a brief acivation in the presence or absence of the ligand, the cells were lysed and the FcRn was immunoprecipated out of the lysates using the anti-FcRn heavy chain antibody, 1G3. The immunoprecipitates were subjected to SDS-PAGE, and the presence of radioactive phosphate was visualized with a phosphoimager plate. Cross-phosphorylation of the FcRn-TrkA chimera was observed at the binding pH (pH 6) in the presence of Fc and IgG but not in the absence of the Fc or IgG or at pH 8. Cross-phosphorylation of the FcRn-TrkA chimera was observed at both pH values in the presence of the anti-FcRn heavy chain antibody 1G3, but not in the presence of 1G3 Fab, or a mutant Fc with no FcRn-binding sites (nbFc) at either pH. Cross-phosphorylation of the FcRn-TrkA chimera was also observed at the binding pH in the presence of the hdFc. Because the hdFc has only one FcRn binding site it cannot bridge FcRn molecules. The induction of cross-phosphorylation of the FcRn-TrkA chimera on the cell membrane is therefore consistent only with the hdFc inducing the formation of the FcRn dimer on the membrane (Figure 7). These results were reproduced six times.
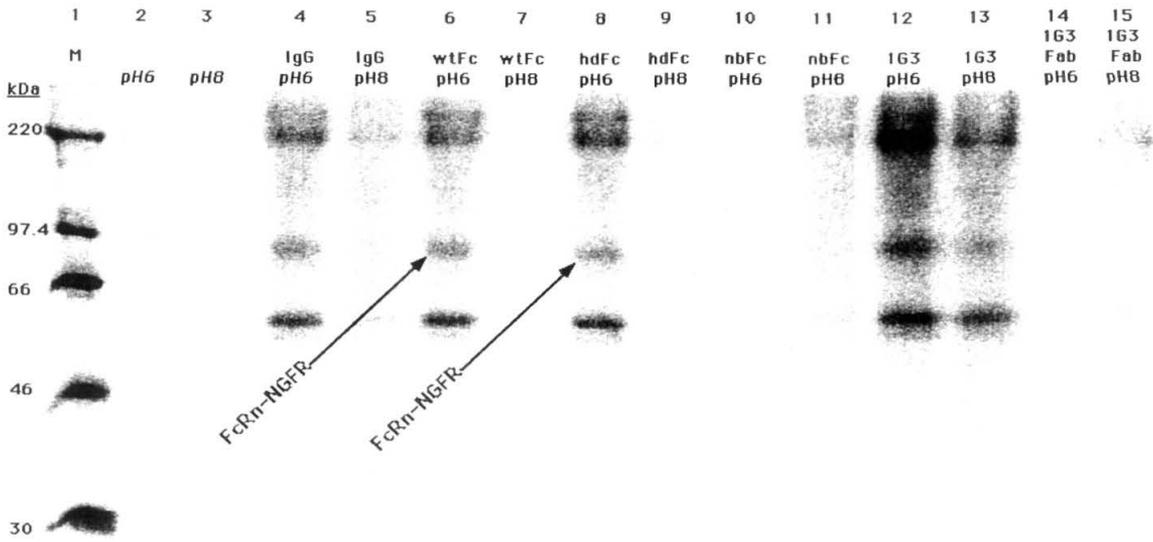
Figure 6: Ligand-induced cross-phosphorylation of FcRn-TrkA (FcRn-NGFR) chimeras by IgG, wtFc, and hdFc. Cells expressing the chimera are briefly exposed to the ligand. Cells are lysed, and lysates are immunoprecipitated with 1G3, an anti-FcRn antibody. Immunoprecipitates are analyzed by SDS-PAGE and the presence of ^{32}P is visualized with a phosphoimager plate.
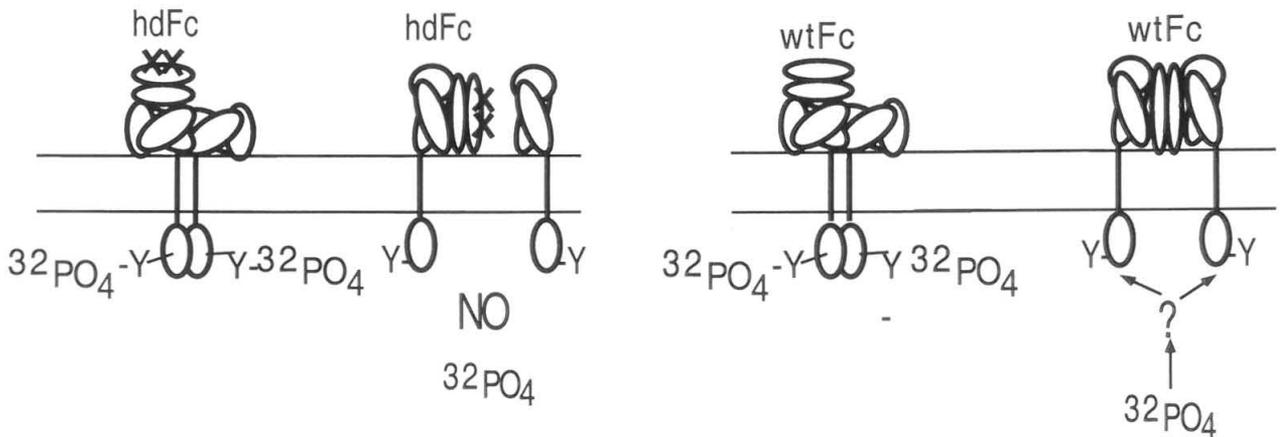
Figure 7: Cross-phosphorylation in the presence of the heterodimeric Fc (hdFc) is diagnostic for the presence of the FcRn-dimer. Cross phosphorylation in the presence of the wild-type Fc (wtFc) does not necessarily indicate the presence of the FcRn dimer.

To determine if the hdFc-induced cross-phosphorylation of FcRn-TrkA could be disrupted by perturbing the FcRn dimer interface, a series of FcRn-TrkA chimeras were constructed with mutations in the FcRn ectodomain. These mutants, one bearing the substitutions His250Glu and His251Glu and the other Asn192Glu, Ser193Glu, Asn225Asp, Cys226Ser, His250Glu, and His251Glu were transfected into CHO-K1 cells and stable clones were isolated as above. When these cells were assayed for induction of cross-phosphorylation with each of the ligands tested above, neither cell line showed any induced cross-phosphorylation. However, at this time the cell lines expressing the original FcRn-TrkA chimera with the wild-type FcRn ectodomain showed no induction of cross-phosphorylation either. In these experiments all the bands in all the lanes were the same intensity, though the intensity of the bands varied between experiments (Figure 8). This experiment was repeated 13 times. To further probe the failure to reproduce ligand-induced cross-phosphorylation, the assay was repeated but the samples were analyzed to detect phosphorylation of tyrosine specifically instead of the presence of radioactive phosphate. The samples were generated as before but split into two gels for two blots. One blot, probed with 1G3 (Figure 9), shows the presence of the FcRn heavy chain in all the lanes. The other blot (Figure 9), probed with a cocktail of anti-phosphotyrosine horseradish peroxidase antibody-conjugates (Zymed) showed no induced bands in any lanes.

Cells transfected with FcRn TrkA
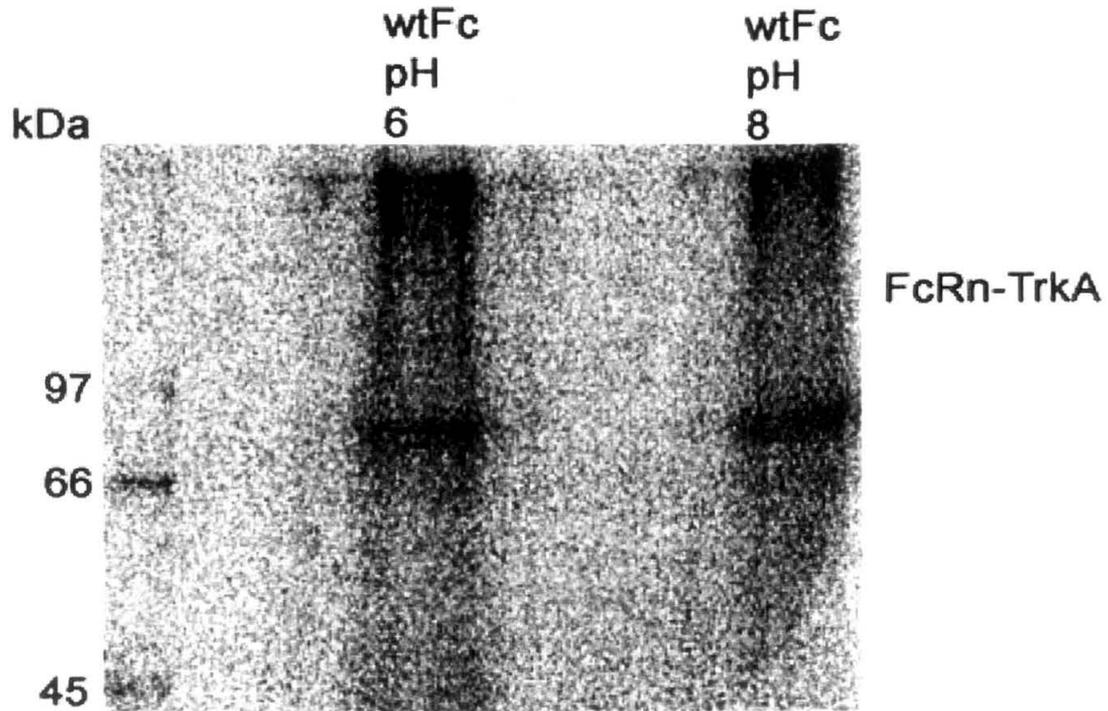
Figure 8: Constituitive cross-phosphorylation of FcRn-TrkA chimeras. Cells expressing the chimera are briefly exposed to the lidand. Cells are lysed, and lysates are immunoprecipitated with an anti-FcRn antibody. Immunoprecipitates are analyzed by SDS-PAGE and the presence of ^{32}P is visualized with a phosphoimager plate.

Blot probed with CA113[67] anti-FcRn anti-sera

Cells transfected with GPI-linked FcRn					Cells transfected with FcRn-TrkA					
serum starve					serum starve					
wtFc	wtFc	wtFc	wtFc	No	wtFc	wtFc	wtFc	wtFc	No	
pH 8	pH 6	pH 8	pH 6	Acti-vation	pH 8	pH 6	pH 8	pH 6	Acti-vation	220
										66 ← FcRn-TrkA
										kDa
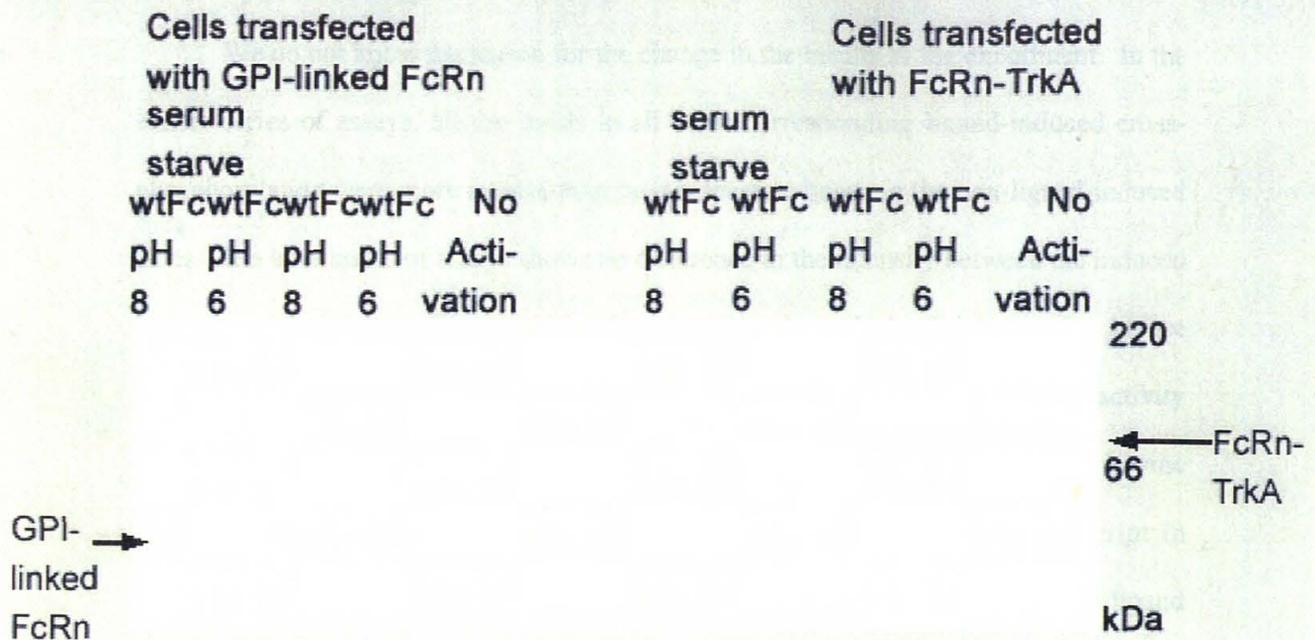
GPI-linked FcRn →

Blot probed with anti-phosphotyrosine cocktail

Cells transfected with GPI-linked FcRn					Cells transfected with FcRn-TrkA					
serum starve					serum starve					
wtFc	wtFc	wtFc	wtFc	No	wtFc	wtFc	wtFc	wtFc	No	
pH 8	pH 6	pH 8	pH 6	Acti-vation	pH 8	pH 6	pH 8	pH 6	Acti-vation	220
										66 ← FcRn-TrkA?
										kDa
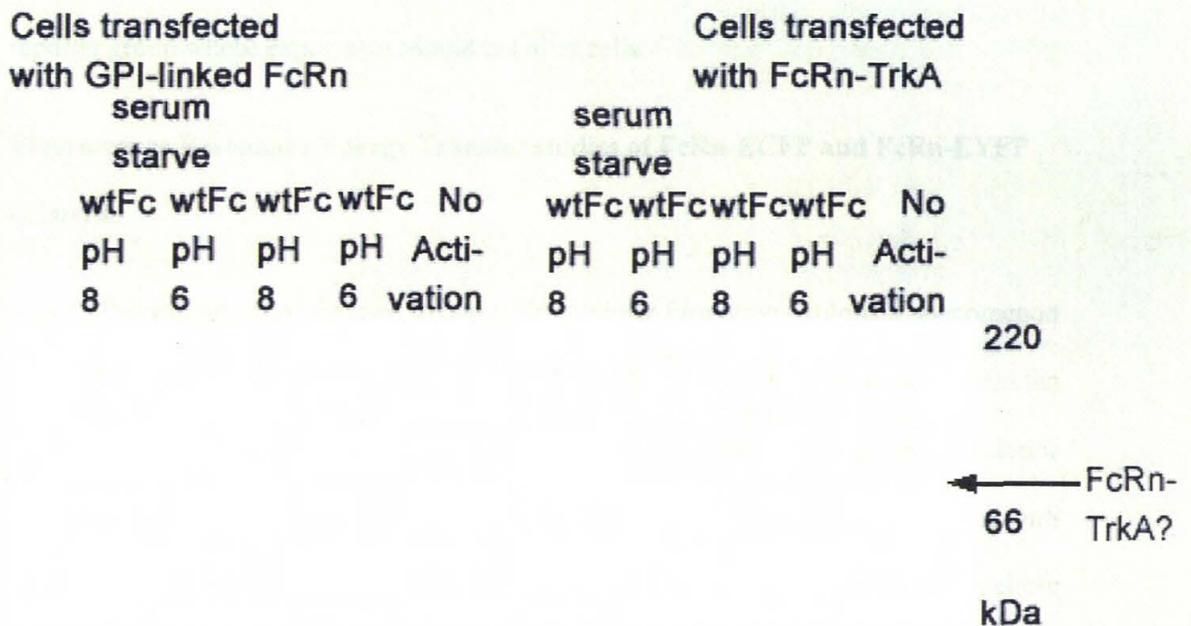
Figure 9: Western Blots of FcRn-TrkA show no ligand-induced cross-phosphorylation. Cells expressing FcRn-TrkA or a GPI-linked form of FcRn were exposed to the ligand, lysed and immunoprecipitated with 1G3, an anti-FcRn antibody. Immunoprecipitates were subjected to SDS-PAGE and Western blot transfer. PVDF membranes were immuno-blotted with anti-FcRn rabbit anti-sera or anti-phosphotyrosine cocktail. Antibody presence is detected with horse radish peroxidase.

We do not know the reason for the change in the results of the experiment. In the earlier series of assays, all the bands in all lanes corresponding ligand-induced cross-phosphorylation were more intense than their respective bands in the non-ligand-induced lanes. The later series of assays shows no difference in the intensity between the induced and noninduced lanes. We speculate that signaling through the TrkA tyrosine kinase domain may have altered the cells at some point such that the phosphorylation activity became constituitive. Recent experiments in our lab have demonstrated that the bovine IgG in serum binds to FcRn in cultured cells (Ramalingham et al., manuscript in preparation). It might be that extended growth in the presence of an activating ligand subjected the cells to an unanticipated selection pressure. Because of problems with reproducibility in this assay, we designed another assay that involves a non-enzymatic reporter group whose expression should not alter cells.

Fluorescence Resonance Energy Transfer studies of FcRn-ECFP and FcRn-EYFP chimeras

The second assay we used to detect FcRn dimer formation exploits a phenomenon called fluorescence resonance energy transfer (FRET). This is an interaction between the electronic excited states of two chromophore molecules. One chromophore, the donor (enhanced cyan fluorescent protein; ECFP), has an emission spectrum that overlaps with the excitation spectrum of the other chromophore, the acceptor (enhanced yellow fluorescent protein; EYFP) (Table II). Excitation is transferred from a donor molecule to an acceptor molecule without the emission of a photon. FRET is most useful when the emission spectrum of the acceptor molecule may be readily distinguished from the

emission spectra of the donor. In this case one may detect FRET by exciting at the donor excitation wavelength and detecting FRET in the form of emission at the acceptor emission wavelength. In the absence of FRET, excitation at the donor excitation wavelength will result in emission at the donor emission wavelength. FRET is useful for studying the association of biological macromolecules because it is dependent on the inverse sixth power of the intermolecular separation (Stryer and Haugland, 1967). Therefore, even with the correct donor and acceptor molecules FRET only occurs when they are close to each other (10-100 Å).

Table 2

Excitation and emission wavelengths

Chromophore	Chromophore single-photon λ of excitation maxima	Single-photon excitation λ from microscope laser	Chromophore two-photon λ of excitation maxima	Two-photon excitation λ from microscope laser	Chromophore single-photon λ of emission maxima
ECFP	425, 450	458	850	800	475
EYFP	515	488	1050	800	535

To exploit FRET in studying FcRn, chromophores were attached to the receptor molecules. Taking advantage of the development of enhanced cyan and yellow versions of green fluorescent protein (GFP), FcRn-ECFP and FcRn-EYFP chimeras were constructed. These chimeras had the fluorescent proteins fused at the carboxy terminus of the complete FcRn heavy chain, such that the propensities of the ecto-, transmembrane, and cytoplasmic domains to form a dimer could be examined collectively.

The FcRn-ECFP and EYFP chimeras were co-transfected into Cos-7 cells and expression of the receptor was detected by the presence of the chromophore. Cell-surface expression of the receptor was detected by live-cell staining with the anti-FcRn antibody 1G3 (Figure 10) but the majority of the fluorescence from ECFP and EYFP was found in intracellular compartments. Because the cells were transfected with both FcRn-chromophore chimeras three possible dimer species occur on the cell surface, FcRn-ECFP/FcRn-ECFP, FcRn-ECFP/FcRn-EYFP, and FcRn-EYFP/FcRn-EYFP. Assuming the association of the chimeras is random, the three possible dimer species occur 25%, 50%, and 25% of the time respectively. Only the second species could give rise to a FRET signal we are detecting, the other two species contribute only to the noise.

In order to calibrate the amount of FRET expected from dimer formation, we constructed positive and negative dimer control chimeras. To make a positive dimer control we needed an obligate heterodimer, we chose the MHC class II molecule HLA-DR. This type I membrane protein has α and β chains to which we fused ECFP and EYFP at the carboxy-termini of the cytoplasmic tails. In transfected cells, only the HLA-DRα–ECFP/HLA-DRβ–EYFP chimera would form and the chromophores would necessarily be close enough together to undergo FRET. As a negative control we needed a molecule that had no known tendency to associate with FcRn. We chose an MHC class I related protein, HFE. In transfected cells, any FRET observed would not be the result of any specific interaction between FcRn-ECFP and HFE-EYFP .

In collaboration with Dr. Mary Dickinson of Scott Fraser's lab we accomplished detection of the FRET signal using a Zeiss LM-510 two-photon laser-scanning confocal fluorescence microscope in its two-photon excitation mode. A chromophore molecule
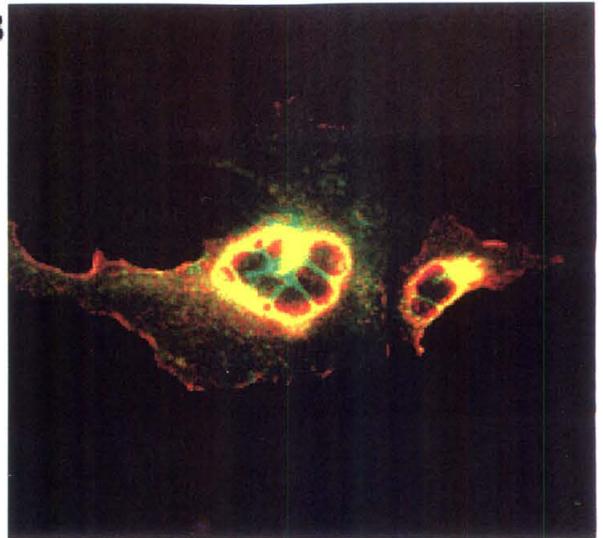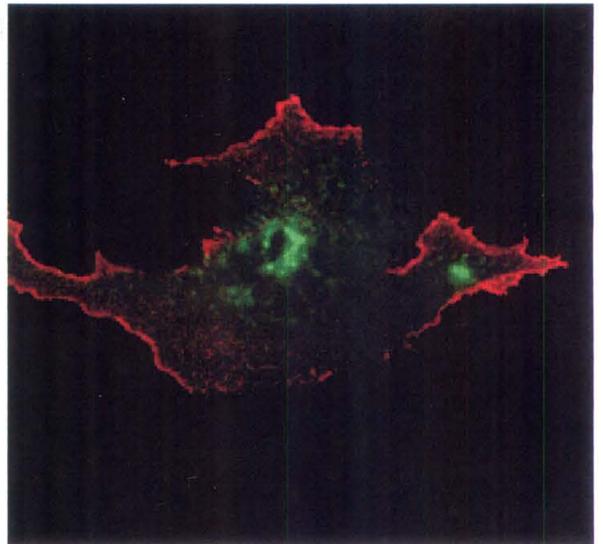
Figure 10: Images A-D are sections from the top -> down through the same cell. A) 1.0 mm into cell; B) 5.0 mm into cell; C) 7.0 mm into cell; D) 10.0 mm into cell. FcRn-EGFP is expressed on the cell surface of transiently transfected COS-7 cells. Live cells were stained with an anti-FcRn antibody, 1G3 and with a goat anti-mouse Alexa-Fluor 568 nm conjugate. Green fluoresence is intracellular FcRn-EGFP. Red fluorescence is the Alexa-Fluor 568 nm conjugate. Yellow fluorescence is the colocalization of the green and red. We argue that the red and yellow fluorescence are indiciative of cell surface FcRn.

may be excited by a single photon if it has the correct energy. Two-photon excitation of a chromophore occurs when two photons, each with half the energy (twice the wavelength) required for a single photon, bombard the chromophore within one femtosecond (Hell et al., 1995). Two-photon excitation is a rare event that requires a large flux of photons to generate excitation at detectable levels but it has several advantages generally over confocal microscopy and one particular advantage for this experiment. Compared to confocal microscopy, two-photon microscopy has a higher signal to noise ratio because it only excites chromophores in the focal point of the sample. Therefore, the whole sample is less photobleached by the excitation and there is almost no light leakage from areas in the sample that are not in the focal plane. In this experiment, two-photon excitation is particularly useful because the separation of the two-photon excitation maxima is twice the separation of the single photon maxima (see Table II). Thus there is a greater chance of exciting only the cyan chromophore with the laser. At the single photon excitation wavelengths it is very difficult to excite the cyan fluorescent protein without also exciting the yellow fluorescent proteins in a FRET-independent manner.

We excited the samples with a 50 kW, 76 MHz, 200 fs pulse of 800 nm light from a titanium-sapphire laser. The light emitted from the sample in this microscope is sent to directly to a spectrophotometer so that the spectra of the different samples, not solely the total intensity, may be compared. When we excite cells transfected with FcRn-ECFP we see emission spectra that peak at the cyan emission wavelength peak (Figure 11b). When we excite cells transfected with FcRn-EYFP with this 800 nm pulse we see no light emitted from all but the very brightest samples. When we excite the cells transfected
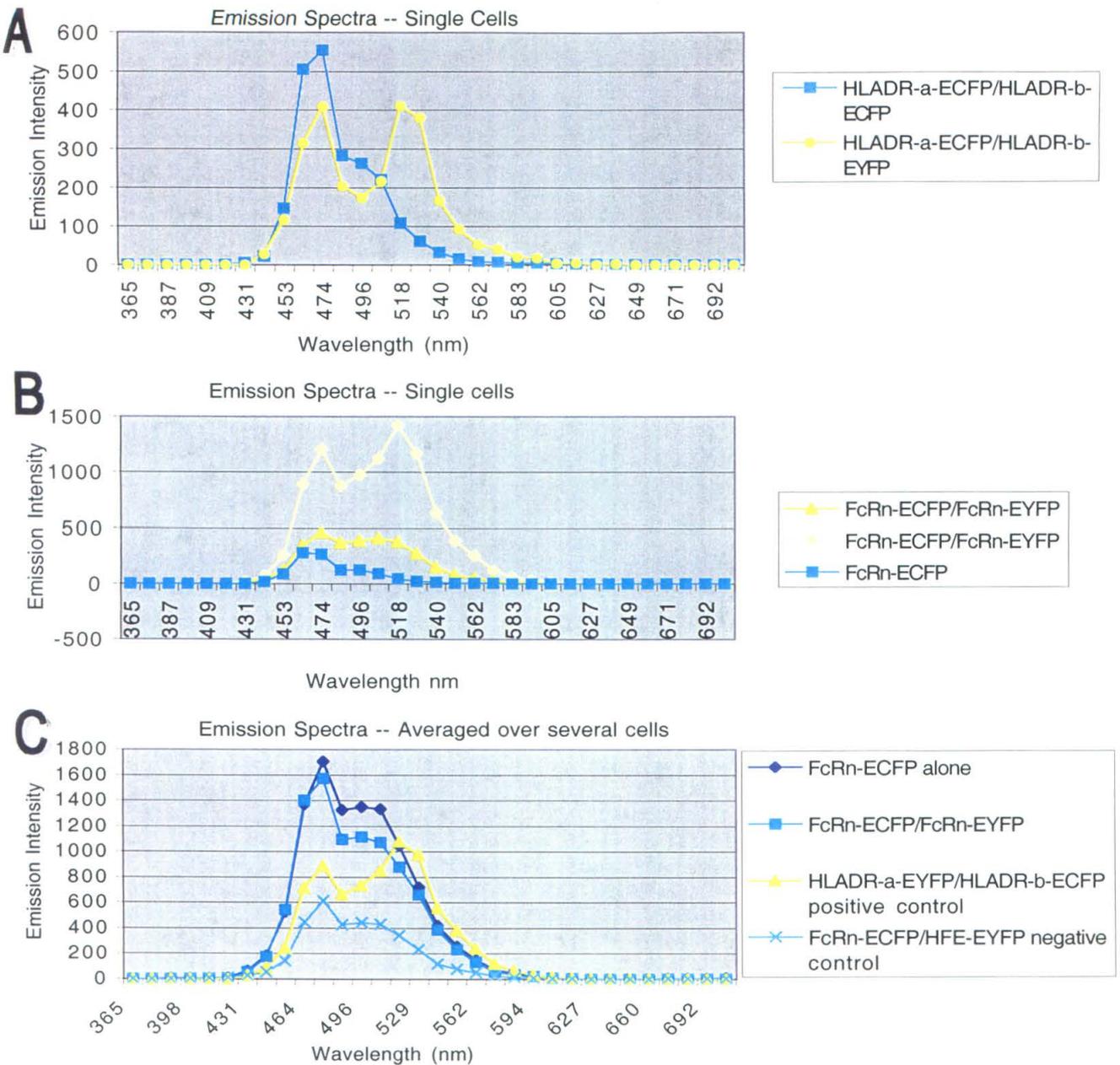
Figure 11: Emission Spectra from the two-photon (800 nm) excitation of cells transfected with ECFP and EYFP chimeras. (A) and (C) Cells transfected with HLADRα-EYFP and HLADRβ-ECFP exhibit FRET. (B) Some cells transfected with FcRn-ECFP and FcRn-EYFP exhibit FRET. (C) Some cells transfected with FcRn-ECFP and FcRn-EYFP or with FcRn-ECFP and HFE-EYFP do not exhibit FRET.

with both FcRn-ECFP and FcRn-EYFP, we see emission at both the cyan and yellow wavelengths in some of the cells (Figure 11b). If the cells are not extremely bright, emission at the yellow wavelength under these conditions is almost certainly the result of excitation by the cyan emission (FRET), and not by the two photon light. Some of the cells transfected with both FcRn-ECFP and FcRn-EYFP do not exhibit any emission peak at the yellow emission (535 nm, Figure 11c). As a positive control series, we also transfected cells with obligate heterodimer cyan and yellow fluorescent protein chimeras. Two-photon excitation of cells transfected with HLA-DRα ECFP and HLA-DRβ ECFP generated an emission spectra peaked at the cyan emission maximum (Figure 11a). Two-photon excitation of cells transfected with HLA-DRα EYFP and HLA-DRβ EYFP resulted in no emitted no light. When cells transfected with HLA-DRα ECFP and HLA-DRβ EYFP were excited by the 800 nm pulse, there was emission of light at both the cyan and yellow fluorescent protein maxima (Figure 11a and 11c). As a negative control, cells were transfected with FcRn-ECFP and a hereditary hemochromatosis linked MHC class I homolog molecule (HFE)-EYFP chimera. When excited with an 800 nm pulse of light, these cells generated fluorescence spectra with little emission at the yellow emission maxima (Figure11c).

While the occurrence of FRET in some cells transfected with FcRn-ECFP and FcRn-EYFP is consistent with the chromophores being within 100 Å of each other it is not necessarily the case that these molecules have formed the FcRn dimer seen in the crystals. An alternative explanation is that the FcRn expression level is so high in the transfected cells that there are molecules within 100 Å of each other at expression bottlenecks inside the cell. Another difficulty in interpreting these results is explaining

why some cells transfected with FcRn-ECFP and FcRn-EYFP exhibit FRET while other such cells do not. The cells that exhibit FRET have a characteristic fluoresence pattern when viewed with epifluorescence, but whether that pattern is indicative of the biologically relevant expression is unknown. Also, as seen in Figure 10, though there is expression of the FcRn chimeras on the surface of the cell, the bulk of the expression is intracellular, in the perinuclear region. This expression pattern is consistent with cells that are over-expressing a protein.

Most of these difficulties can be addressed by generating a stably transfected cell line, preferably with polarized cells, similar to cells that express FcRn *in vivo*. T.S. Ramalingham in our lab has generated and characterized the expression of FcRn-EGFP in Mabin-Darby canine kidney (MDCK) cells. These polarized cells, like the cells of the vascular endothelia where FcRn is expressed, grow in tight monolayers and present a cellular architecture that is convenient for addressing several aspects of FcRn mediated IgG transport. The MDCK cells transfected with FcRn-EGFP express the protein at physiological levels and these cells exhibit the biological function of FcRn, i.e., they trancystose IgG. The locus of the expression of FcRn in MDCK cells, in apical endosomes, at the apical cell-surface, or at the basolateral cell-surface may be modulated by the presence of IgG. Isolating stable MDCK cell transfectants expressing the FcRn-ECFP and FcRn-EYFP chimeras would present a homogenous cellular population for FRET studies. Furthermore, as FcRn in MDCK cells exhibits the biological function of the receptor, we might be able to address whether the chimeras exhibit FRET while trancytosing IgG. As we can modify the locus of FcRn expression in these cells we can attempt to examine the incidence of FRET as a function of cellular localization.

Conclusions

The experiments described here are attempts to determine if the FcRn dimer observed in some crystal forms is physiologically relevant. Analyses of the contacts in the crystal structures are equivocal in their characterization of the contact. The sedimentation equilibrium and dynamic light scattering experiments undertaken by our collaborator clearly indicate that the soluble ectodomain of FcRn is a monomer in solution at concentrations as high as one millimolar. There have been two analyses of FcRn chimaeras attempting to characterize the behavior of FcRn in a membrane. In the first study, FcRn-TrkA chimeras were used to show ligand-induced cross-phosphorylation but these results can no longer be reproduced perhaps because of an alteration of the FcRn-TrkA cell-line. In the second study, cells transfected with FcRn-ECFP and FcRn-EYFP chimeras exhibited FRET under some conditions but is difficult to assess if that sigal is do to the formation of the FcRn dimer. Some of the ambiguities in interpreting the FRET experiments may be alleviated by the generation and characterization of stable transfectants expressing the FcRn-ECFP and FcRn-EYFP chimeras.

Materials and Methods

Assessing the biological significance of a crystal contact by the amount of buried solvent accessible surface area: To calculate the probability of finding an crystal contact that buries greater than 700 Å^2 in size in a monomeric crystal, we employed the formula of Janin:

$$P(B) \sim 4.2 \exp (-B/260) \qquad [1]$$

where B is the amount of surface area buried at the contact on both proteins. Using this formula we can say what the probability of finding an interface of the size found in the FcRn crystals (1780 Å) would be if the FcRn crystals were crystals of monomeric proteins with one molecule in the asymmetric unit.

Assessing the biological significance of a crystal contact by the surfaces' complementarity: To determine the shape complementarity of the FcRn-FcRn dimer interface, we used the "sc" component of the CCP4 suite. Developed by Lawrence and Colman, this metric finds the median of the population of complementarity values, a score of 1.0 is perfectly complementary, a score of 0.0 means the surfaces show not complementary. These values are computed between points on the molecular surface of the interface sampled at 15 dots per square angstrom. The complementarity value is calculated between the point on one surface and the point on the other surface closest to it. It is calculated using the distance between the two points and the unit vectors extending normal to the surface from the first point to the opposite surface and normal to the surface from the second point extending inward away from the first point. The value is computed multiplying the dot product of the unit vectors and an exponential term where the exponent is the product of a weighting term and the square of the difference between the two points. The overall statistic is the average of the medians from the populations of each surface with respect to the other. Derived in this way the shape complementarity statistic has two advantages which are germane to our review. At short inter-surface distances the complementarity of the shape dominates the term while the distance dominates a larger distances. Further, because the average of the median complementarity of each surface for the other is taken, the metric is insensitive to areas of the interface which are not complementary and only weights those areas that are truly in contact with one another.

Assessing the biological significance of a crystal contact by comparing the conservation of the sequence at the surface of the contact: Elcock and McCammon (2001) begin their definition of sequence entropy by adopting the convention for describing sequence entropy employed by the Homology Derived Secondary Structure of Proteins database (HSSP) generated by the European Molecular Biology Laboratory at Heidelberg (ftp://ftpembl-heidelberg.de/put/databases/hssp; Sander and Schneider, 1993). This convention describes the sequence entropy at position i in an alignment s(i) as

$$s(i) = (-1) \Sigma \, p_k * ln(p_k), \qquad\qquad [2]$$

where p_k is the observed probability of finding a given residue k at position i. As may be seen this definition weights all amino acid substitutions equally. To weight substitutions in a manner more indicative of the degree of chemical change caused by the substitution, the amino acids are classed according to Mirny and Shakhnovich (Mirny and Shakhnovich, 1999). Briefly the amino acids are classed as follows: (1) Arg, Lys; (2) Asp, Glu; (3) His, Phe, Trp, Tyr, Val; (4) Asn, Gln, Ser, Thr; (5) Als, Ile, Leu, Met, Val; (6) Gly, Pro. In order to compare the sequence entropies of given surfaces Elcock and McCammon choose to weight the entropy of a given sequence according to its participation at the surface. This surface sequence entropy is defined as

$$SSE(x) \;\; = \Sigma \, p_k * ln(p_k) * ASA_{(sidechains)} / \Sigma ASA_{(allatoms)}, \qquad [3]$$

where p_i is the probability of finding residue type i at that position, and ASA is the associated surface area. Main chain atoms are not included in the numerator of the term, except in the rare cases where a mutation causes a change in the main-chain conformation the main chain atoms have essentially zero sequence entropy. To determine whether a contact surface is biologically meaningful or a nonspecific crystal contact the ratio:

SSE(contact-surface)/SSE(non-contact-surface) is calculated. If this ratio is greater than one the surface entropy of the contact is greater than the rest of the surface and the contact is deemed nonspecific.

Sedimentation equilibrium: These experiments and their written summary was provided by Ghirlando Rudolfo of the NIH. Sedimentation equilibrium experiments were conducted at $4.0^{\circ}C$ on a Beckman Optima XL-A analytical ultracentrifuge. FcRn samples in 0.15M NaCl and 0.063M sodium phosphate (pH = 6.0) were analyzed at different loading concentrations and different rotor speeds as noted in Table 1. Data were acquired as an average of 8 absorbance measurements at a radial spacing of 0.001 cm and nominal wavelengths ranging from 280 to 365 nm depending on the sample loading concentrations (Table 1). Equilibrium, as determined by scans taken six hours apart, was usually achieved within 48 hours. Data were analyzed in terms of a single ideal solute to obtain the buoyant molecular mass, $M(1 - v\rho)$, using the Optima XL-A data analysis software (Beckman) running under Microcal Origin 3.78, by fitting data from each scan to

$$A_r = A_o \exp[HM(1 - v\rho)(r^2 - r^2_o)] + E, \qquad\qquad [5]$$

where A_o is the absorbance at a reference point r_o, A_r is the absorbance at a given radial position r, H represents $\omega^2/2RT$, ω the angular speed in rad s^{-1}, R is the gas constant, T is the absolute temperature and E a small baseline correction. M represents the molecular mass of the glycoprotein, v its partial specific volume and ρ the solvent density. The residuals to the fit were calculated. In all cases, a random distribution of the residuals around zero (± 0.02) was obtained as a function of the radius.

Dynamic light scattering: These experiments and their written summary was provided by Ghirlando Rudolfo of the NIH. The translational diffusion coefficient D was measured from autocorrelation analysis of the quasielastically scattered light. The $\lambda = 514.5$ nm emission of an argon ion laser (Lexel, Model 95) was used in the TEM_{00} mode. The output power was adjusted to 100 mW such that a constant light intensity was obtained. Autocorrelation functions were collected at 22° C using a Brookhaven Instruments BI-9000 AT autocorrelator at an angle θ of 90° with sampling times of 1.0 μs to 100 ms.

100 μL of the sample were transferred to a small volume fluorescence cuvette (Hellma) and autocorrelation functions were accumulated for 5 to 10 minutes. Normalized autocorrelation functions $g_1(t)$ were obtained from the autocorrelated function $G_1(t)$ and the measured baseline, b, as follows:

$$g_1(t) = (G_1(t)/b - 1), \tag{6}$$

The second cumulant of $g_1(t)$, Γ_2, determined using the Brookhaven Instruments analysis software was used to calculate the diffusion coefficient:

$$D = \Gamma_2(4\pi n/\lambda)\sin(\theta/2))^{-2}, \tag{7}$$

where n is the solution refractive index. The equivalent Stokes radius, r, for the FcRn is calculated as follows:

$$r = kT/6\pi\eta D, \tag{8}$$

where k is Boltzmann's constant and η the solvent viscosity.

FcRn concentrations were determined using an extinction coefficient ε at 280 nm of 84,900 M^{-1} cm^{-1}.

Construction of the FcRn-TrkA chimeric construct: Yang Liu generated the first FcRn-TrkA chimera vector used in these experiments. The DNA for the secretion signal sequence and the ectodomain of the mature rat FcRn heavy chain (residues 1-276) and the DNA for the transmembrane and cytoplasmic domains of rat TrkA (residues 417-799) were amplified by PCR. The 5' end of the 3' primer for the FcRn amplification contained the sequence complementary to the 5' end of transmembrane domain of TrkA. The 5' end of the 5' primer for the TrkA amplification contained the sequence complementary to the 3' end of the FcRn ectodomain. The amplified pieces of FcRn and TrkA were purified and then fused by PCR with the 5' FcRn and 3' TrkA primer. The resulting construct fused the residues SPARSS (271-276) from FcRn to TPFGVS (417-422) of TrkA. This construct was subcloned by XhoI – NotI double digestion from pBluescript IISK – (Stratagene) into pBJ-5GS. This vector allows for selection under glutamine free media and amplification in the presence of MSX.

Generation and characterization of FcRn-TrkA stable transfectants: Yang Liu created the CHO-K1 cell line stably expressing the FcRn-TrkA chimera. The vector with the FcRn-TrkA chimera was cotransfected with the vector carrying rat β2 microglobulin into Chinese hamster ovary (CHO-K1) cells. Expression of the FcRn-TrkA chimera on the surface of the cells was confirmed by FACS analysis. FcRn was detected with the anti-FcRn antibody 1G3, the presence of 1G3 was reported by a goat anti-mouse IgG R-phycoerythrin conjugate.

Detection of Fc binding to FcRn-TrkA cells: wtFc, hdFc, and nbFc were coupled to ^{125}I using Iodobeads (Sigma). The specific activities were 4.0, 4.2 and 5.0 x 10^{-5} mCi/pmol Fc, respectively. Three-fold dilution series of eight samples of each of the proteins were

prepared with constant amounts of radiolabeled protein and increasing amounts of unlabeled protein. wtFc was varied from 1 µM to 0.1 nM wtFc each with 0.1 nM iodinated wtFc; hdFc was varied from 6 µM to 3.0 nM hdFc each with 3.0 nM iodinated hdFc; nbFc was varied from 6µM to 3 nM wtFc each with 3.0 nM iodinated nbFc. Prior to the assay the number of FcRn-TrkA molecules/cell (1×10^5) was quantified by incubating batches of 5×10^5 cells with increasing amounts of wtFc until saturation occurred.

Adherent cells were removed from solid support by incubation in PBS pH 7.4 with 4mM EDTA, 0.5% w/v BSA. Cells were quantified using a hematocrit, 5×10^5 were used per reaction. Cells were washed with 50 mM sodium phosphate pH 6.0/150 mM NaCl/0.5% w/v BSA and resuspended in the same buffer. Suspended cells were incubated with the proteins for 20 minutes, and then pelleted. Cells were washed three times in the incubation buffer before counting in a gamma-counter. Data were analyzed by Scatchard analysis and the dissociation constants were in good agreement with those found in surface plasmon resonance assays.

Generation of soluble FcRn, wild-type, heterodimeric, and non-binding Fcs: Please see Chapter 2.

FcRn-TrkA crossphosphorylation assay. Stably transfected cells expressing the FcRn-TrkA chimeric protein were grown out in ten-centimeter dishes. Twelve hours prior to the assay, cells are incubated in media without serum. Three hours prior to the assay the cells are incubated in phosphate free media and ^{32}P phosphoric acid at 250 uCi per plate. To assay the cells the radioactive media is removed and the cells are incubated for one minute with the protein of interest in 5mL of either 50mM Tris-Cl pH 8.0/150mM NaCl

or 50mM MES pH 6.0/150 mM NaCl. The activation buffer is removed and replaced by 1.5 mL of lysis buffer (50mM HEPES pH 7.8, 250mM NaCl, 1% Triton X-100, 10% glycerol, 25mM NaF, 2mM NaVanadate, 1mM PMSF, 2ug/mL each pepstatin, leupeptin, and aprotinin). The cells are immediately placed on crushed dry ice and kept at –80° C for 20 minutes. The cells are thawed at 4° C, transferred to microfuge tubes, and centrifuged at 14,000 rpm in an Eppendorf 5415C centrifuge at 4° C. The supernatents are saved and assayed with the Pierce protein assay. Volumes of supernatant normalized for protein content were immunoprecipitated with an anti-FcRn monoclonal 1G3 and protein G coupled to sepharose at 4° C. Immunoprecipitates were analyzed with SDS-PAGE either phosphorylation was visualized either by exposing the gel to a phospho-imager plate or by western blot analysis. Two Western blots were conducted in parallel following the method of (Sambrook et al., 1989). The primary antibody for the first blot was 1G3. The secondary antibody was goat anti-mouse horse radish peroxidase conjugate. The other blot used cocktail of antiphosphotyrosine antibody-horse radish peroxidase conjugates available from Zymed.

Construction of FcRn-ECFP and FcRn-EYFP expression vectors: The DNA for full-length rat FcRn and rat β2 microglobulin were each amplified by PCR, incorporating a 5' Asp718 site and a 3' HindIII site. The DNA for rat FcRn and EGFP (enhanced green fluorescent protein, codon optimized for mammalian cell expression, Clontech) were also amplified by PCR, incorporating a 5' Asp718 site and a 3' XhoI site and a 5' XhoI site and a 3' HindIII site, respectively. The PCR products were directionally inserted into pBluescript II SK- (Stratagene). The Asp718-XhoI restricted form of rat FcRn was subcloned into the EGFP-containing pBluescript II SK-. This generated a fusion gene

with DNA encoding a secretion signal sequence, the entire FcRn heavy chain, ecto, transmembrane, and cytoplasmic domains, a leucine-glutamate linker, and then EGFP. The DNA for the full-length rat FcRn, rat $\beta 2$ microglobulin, and the FcRn-EGFP fusion gene were directionally subcloned from pBluescript II SK- into a mammalian expression vector pCB6H, (kind gift of Ira Mellman, Yale). The DNA for ECFP and EYFP (enhanced cyan and yellow fluorescent proteins, Clontech) were also amplified by PCR, incorporating a 5' XhoI site and a 3' HindIII site, respectively. The DNA for EGFP was removed from the pCB6H vector containing the FcRn-EGFP fusion protein. The ECFP and EYFP genes were inserted into the remaining backbone. The pCB6H vector places the inserted genes under the hCMV promoter, and provides the gene for resistance to neomycin.

Generation and characterization of transient transfectants: Cells are grown on coverslips in six well plates. The Superfect (Qiagen) transfection reagent and 2µg of DNA per vector were used to transfect cells using the manufacturer's protocol. Each of the FcRn-EGFP, ECFP, EYFP vectors were cotransfected with the rat $\beta 2m$ vector into Cos-7 cells using the Superfect reagent (Qiagen) and the manufacturer's protocol. The expression of the transgenes could be detected by green fluorescence. Expression of FcRn at the cell surface was confirmed by the following protocol. To inhibit endocytosis all steps were conducted with buffers kept at 4° C and cells kept on ice. Live cells were incubated on ice in phosphate buffered saline pH 7.4 with 1% w/v BSA for ten minutes. They were incubated with an anti-FcRn monoclonal antibody 1G3 for 10 minutes. Cells were washed with PBS pH 7.4 and then incubated with the goat-anti mouse Alexa-fluor 568 nm antibody conjugate (Molecular Probes) for ten minutes. The secondary antibody

was removed by washing and the cells were then fixed with 4% paraformaldehyde and methanol. Images of fluorescence were obtained on a LeicaTCS SP confocal mircroscope. Samples were excited at 488 nm and 568 nm wavelengths. The fluorescence was collected in different channels gated such that none of the fluorescence from EGFP bled through to the fluorescence of the Alexa-fluor.

Detection of fluorescence resonance energy transfer in transiently transfected cells: Cells are grown in ten-centimeter dishes and transfected with 8 µg per vector using the Superfect protocol (Qiagen). Cyan and yellow fluorescence in transfected cells was confirmed on the Zeiss LM-510 two-photon laser-scanning confocal fluorescence microscope in confocal mode by exciting the cells in the growth media at 458 nm and 488 nm respectively.

To evaluate FRET a Zeiss LM-510 two-photon laser-scanning confocal fluorescence microscope was used in two-photon mode using a 40x standard objective with a numerical aperature of 1.2. Before collecting the spectra the growth media is removed from the cells, the cells are washed in PBS pH 6.0 and incubated in PBS pH 6.0 with 1% w/v BSA, this buffer is removed and coverslips are placed on the cells in the dish. Spectra were taken for areas in the field of view with fluoresence and in areas with no fluorescence as background control. The background is subtracted from the signal in the figures.

References

Burmeister, W. P., Gastinel, L. N., Simister, N. E., Blum, M. L. and Bjorkman, P. J. (1994a) Crystal structure at 2.2 Å resolution of the MHC-related neonatal Fc receptor. *Nature* 372, 336-343.

Burmeister, W. P., Huber, A. H. and Bjorkman, P. J. (1994b) Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* 372, 379-383.

Carugo, O. and Argos, P. (1997) Protein-protein crystal-packing contacts. *Prot. Sci.* 6, 2261-2263.

Collaborative Computational Project, Number 4.The CCP4 Suite: Programs for Protein Crystallography. (1994) *Acta Cryst.* D50, 760-763 .

Durchschlag, H. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H.-J. Ed.) pp 45 – 128, Springer-Verlag, Berlin.

Elcock, A. H. and McCammon, A. (2001) Identification of protein oligomerization states by analysis of interface conservation. *Proc. Nat. Acad. Sci.USA*, 98, 2990-2994.

Farooqui T, Franklin T, Pearl DK, Yates AJ. (1997) Ganglioside GM1 enhances induction by nerve growth factor of a putative dimer of TrkA. *J Neurochem.* 68, 2348-2355.

Ghirlando, R., Keown, M.B., Mackay, G.A., Lewis, M.S., Unkeless, J.C. and Gould, H.J. (1995) Stoichiometry and thermodynamics of the interaction between the Fc fragment of human IgG1 and its low-affinity receptor Fc gamma RIII. *Biochemistry* 34, 13320-13327.

Hell, S. W., Soukka, J. and Hanninen, P. E. (1995) Two- and multiphoton detection as an imaging mode and means of increasing the resolution in far-field light microscopy: A study based on photon-optics. *Bioimaging* 3, 64-69.

Janin, J. (1997) Specific versus nonspecific contacts in protein crystals. *Nat. Struct. Biol.* 4, 973-978.

Janin, J. and Rodier, F. (1995) Protein-protein interaction at crystal contacts. *Proteins* 23, 580-587.

Lawrence, M.C. and Colman, P. (1993) Shape complementarity at protein-protein interfaces. *J. Mol. Biol.* 234, 946-950.

Martin, W. L. and Bjorkman, P. J. (1999) Characterization of the 2:1 complex between the class I MHC-related Fc receptor and its Fc ligand in solution. *Biochemistry.* 38, 12639-12647.

Martin, W.L., West, Jr., A.P., Gan, L., Bjorkman, P.J. (2001) Crystal structure at 2.8 Å of an FcRn/heterodimeric Fc complex: mechanism of pH-dependent binding. *Molecular Cell*, 7: 867-877

Minton, A.P. (1998) in *Methods in Enzymology, Vol. 295* (Ackers, G.K. and Johnson, M.L. Eds.) pp 127 – 149, Academic Press, San Diego.

Perkins, S.J. (1986). Protein volumes and hydration effects - the calculations of partial specific volumes, neutron-scattering matchpoints and 280-nm absorption-coefficients for proteins and glycoproteins from amino-acid-sequences *Eur. J. Biochem.* 15, 169-180.

Peters, J. W., Lanzilotta, W. N., Lemon, B. J., and Seefeldt, L. C. (1998) X-ray crystal structure of the Fe-only hydrogenase (CpI) from Clostridium pasteurianum to 1.8 angstrom resolution. *Science* 282, 1853-1858.

Raghavan, M., Chen, M.Y., Gastinel, L.N. and Bjorkman, P. J. (1994) Investigation of the interaction between the class I MHC-related Fc receptor and its immunoglobulin ligand. *Immunity*, 1, 303-315.

Rivas, G., Fernandez, J.A. and Minton, A.P. (1999) Direct observation of the self-association of dilute proteins in the presence of inert macromolecules at high concentration via tracer sedimentation equilibrium: Theory, experiment, and biological significance. *Biochemistry* 38, 9379-9388.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (N. Irwin, Ed.) 2nd Ed. pp 18.60-18.75, Cold-Spring Harbor Laboratory Press, Cold-Spring Harbor, New York.

Sánchez, L. M., Penny, D. M. and Bjorkman, P. J. (1999). Stoichiometry of the interaction between the MHC-related Fc receptor and its Fc ligand. *Biochemistry*. 38, 9471-9476.

Sander, C. and Schneider, R. (1993) The HSSP data base of protein structure-sequence alignments. *Nucleic Acids Res.* 21, 3105-3109.

Simister, N. and Rees, A. R. (1985) Isolation and characterization of an Fc receptor from neonatal small intestine. *Eur. J. Immunol.*, 15, 733-738.

Stryer L, Haugland RP. (1967) Energy transfer: a spectroscopic ruler. *Proc. Natl. Acad. Sci. USA* 58, 719-726.

Vaughn, D. E. and Bjorkman, P. J. (1997) High-affinity binding of the neonatal Fc receptor to its IgG ligand requires receptor immobilization. *Biochemistry*, 36, 9374-9380.

Vaughn, D. E. and Bjorkman, P. J. (1998) Structural basis of pH-dependent antibody binding by the neonatal Fc receptor. *Structure* 6, 63-73.

Vaughn, D. E., Milburn, C. M., Penny, D. M., Martin, W. L., Johnson, J. L. and Bjorkman, P. J. (1997) Identification of critical IgG binding epitopes on the neonatal Fc receptor. *J. Mol. Biol.* 274, 597-607.

West, A. P. and Bjorkman, P. J. (2000) Crystal structure and IgG binding properties of the human MHC-related Fc receptor. *Biochemistry* 39, 9698-9708.

Chapter 5:

A comparision of 171 nonobligate

protein-protein interfaces

This chapter is an analysis of a collection of 171 nonobligate protein-protein interfaces obtained from the protein data bank. Aspects of interfaces quantified in earlier studies are extended here over this larger database. A principal component analysis of the quantifiable aspects of protein interfaces reveals that the interfaces of antibodies and proteases are special cases somewhat distinct from the general population of nonobligate protein-protein interfaces. David Mathog suggested several years ago that I pursue conducting a principal component analysis on interface data. Ben Bornstein of the Machine Learning Group at JPL was a patient mentor both in explaining (repeatedly) the principal component analysis and (with Barbara Engelhardt of the Artificial Intelligence Group at JPL) in writing code that made the completing the patch-wise analyses of this study a trivial exercise.

Introduction

This chapter is a review of a large number of structures of nonobligate protein-protein complexes. The purpose of this chapter is to provide a context for understanding the crystal structure of the nonobligate complex between FcRn and hdFc. The understanding of protein structure has been enriched by analyses of collections of crystal structures. It is hoped that this analysis will enrich the understanding of nonobligate protein-protein complexes and of the FcRn/Fc complex in particular.

Until recently the amount of data available limited the progress in studying nonobligate protein-protein complexes. While still a small number of samples from a statistical point of view, the 171 of samples in this review generated more complete Gaussian distributions for the metrics used than previous studies. This greater completeness made pursuing a principal component analysis of these data more reasonable.

Selection criteria for inclusion in this study

In reviewing nonobligate protein-protein complexes, we select only those complexes solved by X-ray diffraction, at resolutions between 4 Å to 1.2 Å resolution. All complexes between unique combinations of proteins are included. Included in this study are complexes between more than two proteins, each of which is stable on its own. Excluded from consideration are obligate complexes, homo-oligomers, structures solved by NMR, theoretical models, and multiple versions of a complex between identical proteins. This selection of complexes is drawn from the February 15, 2001 release of the Protein Data Bank.

The selection criteria generate a population of complexes shaped by several biases. A selection of crystal structures is biased first towards those proteins interesting to crystallographers and then towards those complexes that crystallize. As a result of the first bias, several types of complexes, antibody/antigen, protease/protein protease inhibitor, major histocompatability complex class I and class II/T-cell receptor, IgG-Fc/IgG-Fc binding proteins, and cell-surface receptors/protein ligands, are more thoroughly represented. The second bias skews the sample towards complexes between small, monomeric proteins, though this bias is diminishing as the Protein Data Bank grows. Finally, because the annotation of the files in the Protein Data Bank is not constructed to process these sorts of selection criteria, the population was assembled by hand. Therefore to be included in the population, the unique nonobligate complex of proteins had to be one that the author could recognize. This unintended selection criterion will bias the sample away from including protein complexes whose nonobligate *Nature* is unknown to the author. These biases, while unfortunate, are unavoidable given the scarcity of the available data and are common to all of the analyses of this type (Jones and Thornton 1996; LoConte et al., 1999).

The categories of nonobligate complexes

In this review we will compare structures 163 cocrystal complex structures containing 171 interfaces. These structures will be into divided into nine categories:

1) 43 interfaces between proteases with protease inhibitors.

2) 32 interfaces between antibodies with their antigens.

3) 17 enzymes complexed with activators, inhibitors or other enzymes.

4) 11 interfaces in which one of the two proteins is >90,000 Daltons.

5) 27 interfaces between G-proteins, cell-cycle proteins, or signal transduction proteins.

6) 11 interfaces in which one of the two proteins is an MHC class I or II molecule or homolog.

7) 5 interfaces between the Fc portion of IgG and another molecule.

8) 19 interfaces between ectodomains of cell surface receptors with their ligands.

9) 4 miscellaneous interfaces.

The categories overlap in some cases, and the categories are determined by different criteria. Further, some interfaces could be included in more than one category. The protease/protease inhibitor and antibody/antigen categories have the least amount of variation. The enzyme complex and G-protein categories are the most heterogeneous populations of structures. The large complexes are separated out from other categories because it is believed that these proteins are more likely to undergo extensive conformational changes upon complex formation (Lo Conte et al., 1999). The MHC class I and class II and homolog complexes are grouped separately because they are a more homogeneous subset of the cell-surface receptor ectodomain category. The Fc structures are separated for the same reason. This loose categorization on biological and structural criteria is adopted as a substitute for a more quantitative index based on the affinity of the interaction. At the moment, affinities of protein-protein interactions are not determined in any standard way nor are they stored in any convenient database. In the absence of such data we separate the interfaces upon the qualitative hierarchy.

The analyses

This review will analyze these interfaces with a combination of metrics readily available to a protein crystallographer. All the calculations except those involving

Voronoi volumes were completed only over the amino acid atoms of the polypeptide chains, sugars, cofactors and ordered waters were omitted. Voronoi volume calculations define atomic volumes by the surrounding atoms and therefore require using all atoms in the pdb file. The chains chosen to represent the given proteins in the given calculations are included in Appendix B. The following types of analyses, defined below, will be conducted on these complexes: interface surface area, circularity, planarity, atom accessibility, atom burial, atom packing, shape complementarity, number of hydrogen bonds, chemical *Nature* of constituent atoms, and residue composition of the interfaces. The numerical results of these analyses are included in Appendix B.

The proteins' masses, volumes, and total solvent accessible surface areas are included as points of reference. The mass of the protein given is the mass of the protein atoms in the crystal. This value, while not necessarily the actual biological value, because regions of the protein that are disordered in the crystal are not included in the coordinates, is close to it. And importantly, it is the value directly related to the surface area values that are calculated with exactly these atoms. The whole protein volume values are calculated with the VOIDOO algorithm (Kleywegt and Jones, 1994) using the default settings and the standard 1.4 Å radius probe sphere. The solvent accessible surface areas are calculated using the algorithm originally written by Lee and Richards, (1971). The chemical group radii are taken from Chothia, (1975) and the radius of the water probe is 1.4 Å.

Protein level analyses

Solvent accessible surface area

The change in solvent associated surface area found upon complex formation is called the interface surface area. The interface surface area is used as a metric for studying protein-protein interactions because there is a correlation between it and the hydrophobic free energy of transfer from a polar environment to a hydrophobic environment (Chothia, 1974). The interface surface area is calculated as in LoConte et al., (1999) on a per atom contribution basis. The total surface for each protein is calculated, as is the total surface for the two proteins in complex. The difference is the interface surface area, and the atoms associated with the surface are known so the amount the amount of interface attributable to given amino acids, chains, and proteins is readily determined.

The mean interface surface area per protein is 1106 Å^2 (standard deviation, 785 Å^2). This is higher than the 983 Å^2 value reported by Jones and Thornton and the 1940 Å^2 per binary complex reported by LoConte et al., (1999). It reflects both the broader sampling of the database in this discussion and the trend as crystallography progresses towards solving larger complexes. The histogram of interface surface areas (Figure 1) reveals a longer version of the tail observed in other studies (LoConte et al., 1999) of larger interface areas for which we have only a few observations. The new data indicate that this mean may continue to trend upward as more data are accumulated.
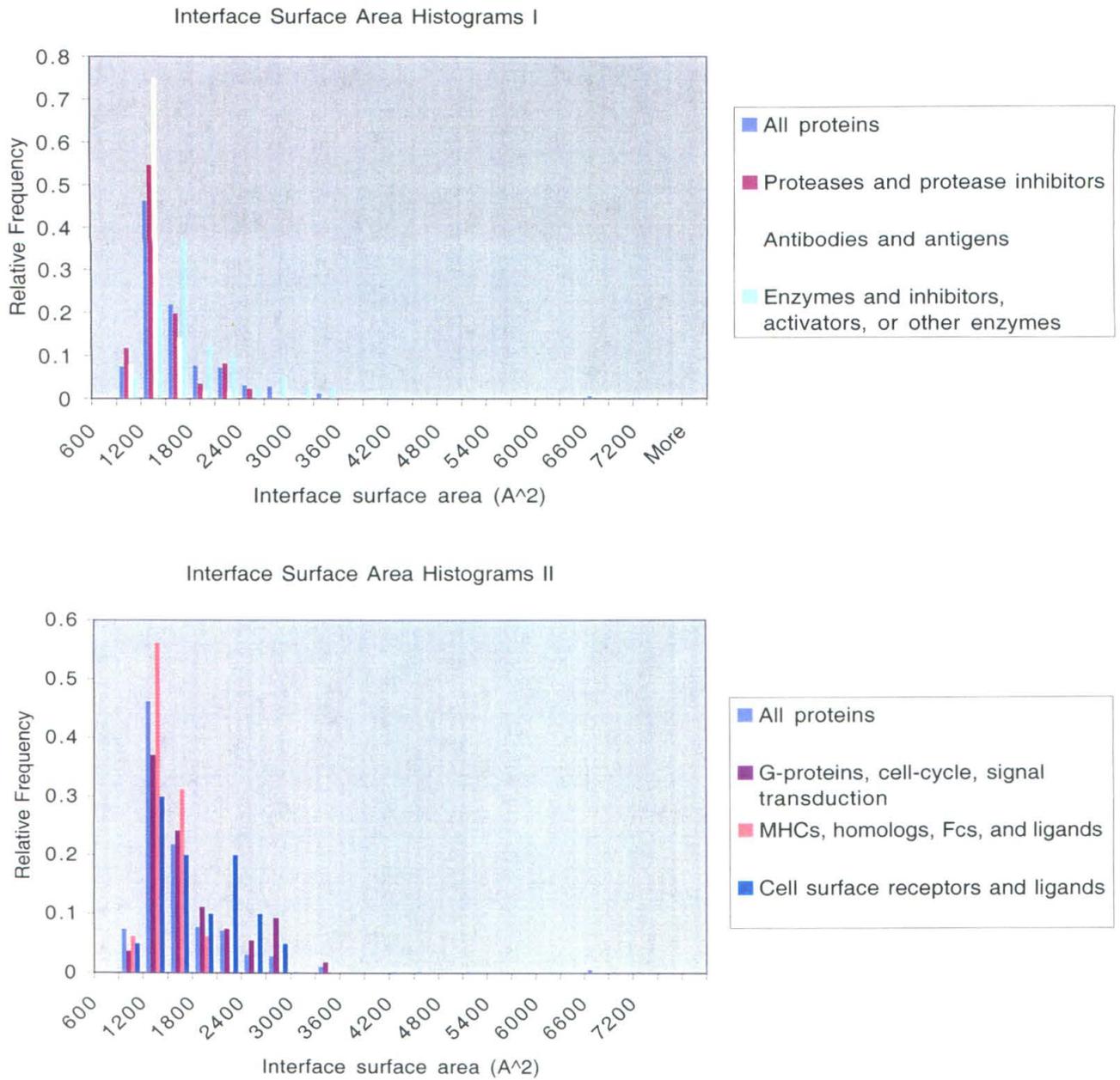
Figure 1: The histograms of interface surface area reveal a tail of larger complexes. This tail extends as more larger structures are solved.

Further support for a continuing upward trend in mean interface surface area may be drawn from the histograms for the different types of protein complexes. Small protease/protease-inhibitor interfaces and antibody/antigen interfaces continue to dominate the data set in terms of frequency and their similar narrow distributions are apparent, peaking at around 1000 Å^2. There is another mode to the protease-protease inhibitor complex, first noted by LoConte et al., which contains the more recently acquired thrombin complexes. Similarly, assuming a normal distribution of interface surface areas, the cell-surface receptor interfaces and the G-protein, cell-cycle, signal transduction interfaces show now that their current peaks may shift higher as more data are assembled. The large complexes are responsible for the largest interface surface areas, and the number of these to be solved is still insignificant. However, the total number of these large complexes employed in biological systems will probably a significantly reduced number. Their effect on the mean may not be overwhelming. It seems reasonable to predict that upon acquisition of most of the known protein-protein interactions that their mean interface area will reach somewhere in the 1200-1300 Å^2 range, rather than the 900-1000 Å^2 range, where it current resides.

Circularity

The circularity of protein interfaces allows assessment of the intimacy of the association between the proteins on a protein-size scale (Jones and Thornton, 1996). The circularity is described by finding the least squares plane through the atoms involved in the interface and taking the ratio of the two principle axes (Jones and Thornton, 1996). In the discussion here, these calculations were conducted using least-squares plane

command of the MOLEMAN2 algorithm from the Uppsala Software Factory (Kleywegt and Jones, 1999). Mathematically, this is arrived at by a principal component analysis of the coordinates, where the eigenvalues of the covariance matrix correspond to the principle axes. By the convention adopted by Jones and Thornton (1996), the second principal component is divided by the first giving a ratio that is always less than 1 but approaching 1 as the interface approaches a perfect circle.

The mean circularity of protein-protein interfaces in this study is 0.43 (standard deviation 0.21). This is significantly less circular than the original work of Jones and Thornton. They found means for circularity ranging between 0.55 and 0.75 for nonobligate interface surfaces. Again the shift is the result of the broader sampling of the data that is now possible. The histograms of interface circularity (Figure 2) reveal that protein-protein interfaces sample the circularity spectrum comprehensively. Antibodies and antigens, the enzyme and inhibitors, activators or other enzyme category, and the graphically combined categories of MHC class I, II molecules, their homologs and ligands, and Fc molecules and ligands are more circular. G-protein, cell-cycle, signal transduction interfaces and cell-surface receptor and ligand interfaces are less so. The extensive sampling of circularity space by nonobligate interface surfaces reveals how crude an approximation is a circular contact between globular proteins.
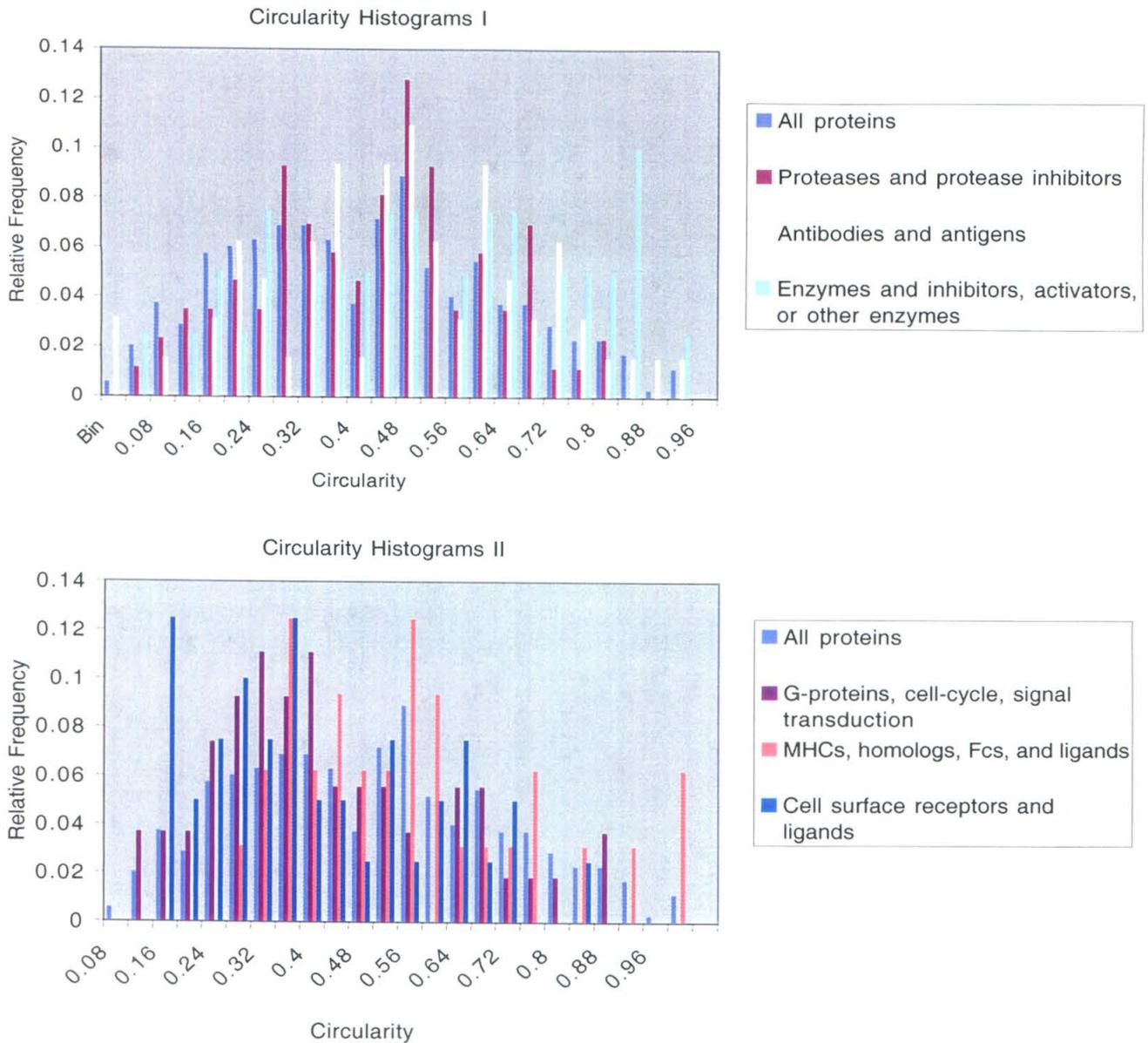
Figure 2: The low, broad distribution in the circularity histograms shows that nonobligate protein-protein complex interfaces assume all manner of oblong shapes.

The planarity

It is intuitively satisfying to consider the third dimension, the thickness of the dispersion about the least-squares plane through the interface, by itself. To compare separate interface surfaces, we use the root mean squared deviation (RMSD) from the plane, to measure the degree of planarity. The RMSD is related to the third (smallest) eigenvalue from the principal component analysis of the dispersion of atoms at the interface. Because this is not a ratio of eigenvalues, we must normalize this third principal component by dividing it by the number of samples used to generate it, by convention we take the square root. Again we used the least squares plane command of MOLEMAN2 to find the principal axes of the least squares plane.

The mean RMSD from the plane of the interface surface in this study is 3.17 Å (standard deviation, 1.36 Å). This is more convoluted than was reported previously (Jones and Thornton, 1996) due to the incorporation of new data. Specifically, we find the antibody/antigen category assuming a monomodal distribution at low RMSD, while the protease/protease inhibitor and G-protein, cell-cycle, signal transduction categories each have monomodal distributions with increasing mean RMSDs (Figure 3). The combined categories of MHC class I, II molecules, their homologs and ligands and Fc molecules and ligands assumes a bimodal distribution. This bimodal distribution is not explained by the combination of these categories as there is considerable overlap in the planarity values for these two groups. This bimodal pattern is adopted also by the enzyme and inhibitor, activator, or other enzyme category as well as the cell-surface receptor and ligand category. However, both of these distributions have higher mean RMSDs than the bimodal immunological category.

Figure 3: The planarity histograms assume slightly bimodal distributions for some of the protein categories.

To determine if the interface surface is more flat than the surface of a protein generally, it is necessary to extract samples of the rest of the surface. Jones and Thornton first attempted this on a smaller data set including 31 heterocomplexes with a patch-wise analysis of the surface (Jones and Thornton, 1997a). Their patch size depended per protein on the size of that protein's interface. They found that in contrast to small molecules the surfaces of proteins most likely to interact with other proteins are more flat than the surface is generally. We endeavored to repeat this analysis on this larger data set. To do this an algorithm had to be developed. We again sampled the surface by taking surface patches but we made their general dimensions constant. This can be done by choosing an atom and constructing a patch of all the atoms on the surface within a given distance. By generating these patches for every surface atom with a patch radius that generates a total patch surface area, equivalent to the average area of a protein-protein interface, we can generate a data set to make an unbiased comparison. A rapid version of this algorithm was written by Ben Bornstein of the Machine Learning Group at JPL.

We find an average RMSD to the plane of surface patches with the average area of 1000 Å^2 to be 4.1 Å (standard deviation 1.5). This is an even greater difference between the average RMSD of interfaces and that of similarly sized surface patches than was previously determined (Jones and Thornton, 1997a). We noted previously how the larger amount of data in this discussion generated a broader range of interface surface areas and a thorough sampling of circularity space extending the descriptive power of these metrics but limiting their use as predictors of what surface is an interface. In the

case of the RMSD from the plane of the interface we find that the addition of data extends the predictive power of this metric.

Summary: protein level analyses

We have considered metrics, which describe protein-protein interfaces on a protein-size scale. The interface surface area describes how much desolvation of protein surface, the hydrophobic effect, is involved in complex formation. The circularity proved to be an evenly sampled characteristic suggesting that interfaces come in all manner of shapes. The RMSD from the plane of the interface is a predictor of what surface of a protein is likely to be involved in an interface. We turn now to different metrics, which assay protein interfaces on the atom-size scale.

Atom level analyses

Atom Accessibility

The first atomic parameter we consider is atom accessibility. It was noted in the mapping of antibody epitopes onto viral coat proteins that the accessible surfaces were those that were usually bound by the antibody (Lou et al., 1987). The atom accessibility metric is a derivative of the residue accessibility metric used in earlier patch-wise analysis of protein surfaces (Jones and Thornton, 1997a). Jones and Thornton found that the average accessibility of the residues of a protein interface had higher average surface area than residues in other patches of surface. Our parameter is defined as the average surface area per interface atom. The surface areas are those calculated by Gerstein's version of the Lee and Richards algorithm (Gerstein, 1992).

The average surface area of an interface atom is 14.4 Å^2 (standard deviation 2.1 Å^2). We find that the atom accessibility on the surface generally 11.1 Å^2 (standard

deviation 1.42 Å^2) is less than atom accessibility at interfaces. This is consistent with the result of the Jones and Thornton (1997). Examining the histograms of the distributions by category (Figure 4) reveals the distributions of the proteases and protease inhibitors and antibodies and antigens categories being shifted below the mean. The other, newer, categories employ even more accessible surfaces of their proteins and their distributions peak above the current mean. Interfaces involve atoms with more surface area than surface atoms have on average and if anything this difference may increase with additional data.

Atom Burial

While compiling the data for the atom accessibility study, we noticed that the amount of surface buried per atom generated a striking series of distributions for our categories (Figure 5). This is distinct from the amount of surface area each atom exposes when the proteins are not in complex because not all atoms bury their entire surface upon complex formation. Proteases and protease inhibitors and antibody and antigen interfaces produce well-separated bimodal distributions with peaks on either side of the mean atom-burial for all proteins. The protease interfaces have lower atom-burial while the inhibitors have higher atom-burial than average. Similarly, the atoms of the antibody interface surface have lower average burial while the antigens have higher than average burial. The enzymes and inhibitors, activators, or other enzymes category distributes broadly about the mean. The other three categories overlay rather tightly with distributions that peak at slightly higher atom surface area buried than the overall average.
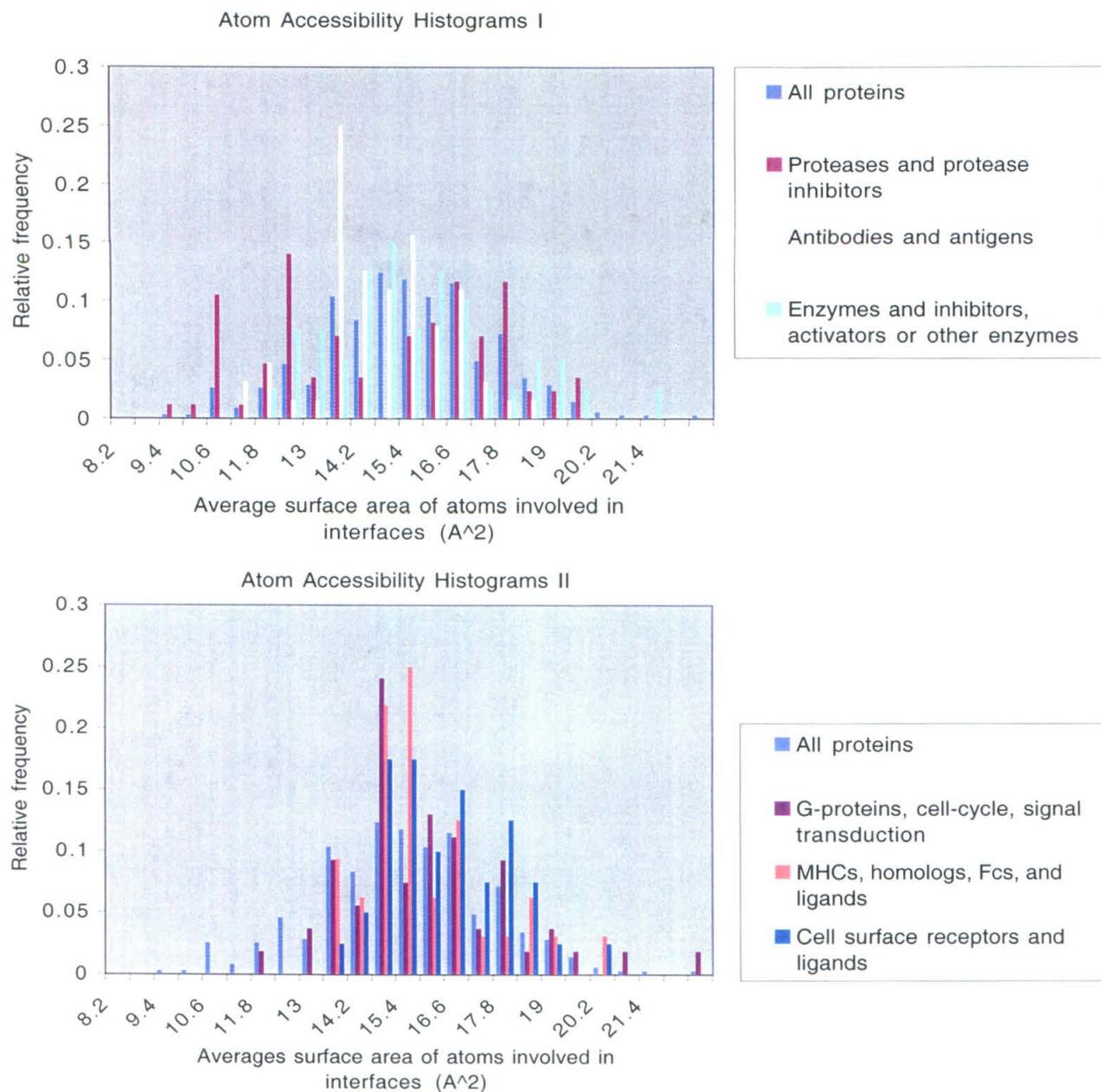
Figure 4: Interface atoms are more accessible than surface atoms generally. These histograms show that the average atom accessibility of proteins solved more recently is even higher suggesting that this difference may increase with additional data.

Atom Burial Histograms I

Legend: All proteins; Proteases and protease inhibitors; Antibodies and antigens; Enzymes and inhibitors, activators, or other complexes

Average surface area per atom buried in the interface (A^2)

Atom Burial Histograms II

Legend: All proteins; G-proteins, cell-cycle, signal transduction; MHCs, homologs, Fcs, and ligands; Cell surface receptors and ligands

Average surface area per atom buried in the interface (A^2)

Figure 5: The atom burial histograms reveal the best alignment with the categories of complexes in this study. Note the bimodal distributions of antibodies and antigens, and proteases and protease inhibitors, while the other categories are monomodal and overlap tightly. This pattern is also seen in the principal component analysis.

Shape Complementarity

To assess how intimately associated two proteins are at an atomic level we turn to the shape complementarity that the two protein interfaces have for each other. By examining directly the distances between and opposition of shapes of opposing surfaces this metric will give us an approximation of the extent and strength of the van der Waals contacts between the proteins. This is the most complex metric in our study, the index of surface complementarity, and it may be computed in a variety of ways. We adopt the "sc," algorithm (Lawrence and Colman, 1993), a part of the CCP4 program suite (CCP4, 1994). We employ it using all the default settings. This metric finds the median of the population of complementarity values. These values may be between 0 and 1, where 1 is perfectly complementary. The complementarity value is calculated between the point on one surface and the point on the other surface closest to it. It is calculated using the distance between the two points and the unit vectors extending normal to the surface from the first point to the opposite surface and normal to the surface from the second point extending inward away from the first point. The value is computed multiplying the dot product of the unit vectors and an exponential term where the exponent is the product of a weighting term and the square of the difference between the two points. The overall statistic is the average of the medians from the populations of each surface with respect to the other. Derived in this way, the shape complementarity statistic has two advantages that are germane to our review. At short intersurface distances the complementarity of the shape dominates the term while the distance dominates a larger distances. Further, because the average of the median complementarity of each surface for the other is taken,

the metric is insensitive to areas of the interface that are not complementary and only weights those areas that are truly in contact with one another.

The mean shape complementarity of the interfaces in this study is 0.69 (standard deviation 0.06). It was previously noted that proteases complexed with protease inhibitors have a higher shape complementarity than antibodies complexed with their antigens (Lawrence and Colman, 1993). Our histograms of shape complementarity (Figure 6) show now that the protease/protease inhibitor and G-protein, cell-cycle, signal transduction categories have higher shape complementarity. Cell-surface receptors and antibodies have similar shape complementarity distributions, both close to the mean for all complexes. The category, enzymes and inhibitors, activators, and other enzymes shows a slightly lower than normal index. This may be an artifact of the omission from these calculations of cofactors and water molecules that are an integral part of some of these complexes. The combined categories of MHC class I, II molecules, homologs, and ligands and Fc molecules and ligands has the lowest average shape complementarity statistic. This is consistent with some crude correlation between shape complementarity and affinity. Six of the 16 complexes in this category contain interfaces between T-cell receptors and major histocompatability molecules, these complexes are selected to have comparatively low (μM) affinity.

Atom Packing

Another method for estimating the importance of van der Waals contacts across an interface is to measure the volume of the atoms at the interface with Voronoi polyhedra (Chothia, 1975; Harpaz et al., 1994). A Voronoi polyhedron is generated around an atom in a protein is generated in two steps. First, lines from that atom to each

Figure 6: The shape complementarity histograms reveal a narrow distribution indicative of the uniform importance of van der Waals forces in protein-protein interactions.

of its neighbors are drawn and second planes normal to these lines are placed according to the van der Waals radii of the two atoms. Each plane extends until it intersects another. Atoms that are not surrounded by sufficient atoms are omitted from the calculations. This method is useful for judging atom packing because, except for a vertex error due to differences in atom radii, it includes all the space around the atoms in the calculation (Gerstein et al., 1995). This calculation was conducted over all the atoms in the protein data bank file, including waters, sugars, metals, and cofactors. These atoms were included in this calculation because of the dependence of this measurement on the neighboring atoms. To assay the packing of atoms at the interface, we separate the interface atoms into two categories, those that are buried completely in the interface, and those that are not. To normalize these values from interface to interface, we divide the average atom volume for both the completely and partially buried interface atoms by the average volume of the atoms buried in the core of their respective proteins. This gives us two ratios that compare the packing of the interface atoms with the extraordinary packing of the protein core (Harpaz et al, 1994; LoConte et al., 1999).

The average volume of atoms completely buried in the interfaces was 1.1 (standard deviation 0.12) times that of the atoms buried in the proteins' cores. This is higher than the 1.03 (standard deviation 0.03) reported earlier (LoConte et al., 1999). The earlier study omitted 11 of 75 complexes from the calculation of this ratio where ours compares all 163 complexes. Despite the larger average volume found in our study for completely buried interface atoms, this volume is small and indicative of close packing. The packing of the atoms partially buried at the interface, is not as dense. The average atom volume of atoms partially buried in the interfaces was 1.9 (standard

deviation 0.26 Å^3) times that of the atoms buried in the proteins' cores. The less dense packing of these atoms may be the result of the omission, in lower resolution structures of what would be ordered water molecules in the interface periphery. The incorporation of such ordered waters in the periphery of interfaces has been proposed (LoConte et al., 1999) as an explanation of the observation that amino acid substitutions in the interface periphery generally have small effects on complex formation (Cunningham and Wells, 1993; Clackson and Wells, 1995). An alternative explanation is that the periphery of an interface is not as well packed as the core of the interface and that the looser packing itself reduces the importance of the amino acids in the periphery to complex formation. Because the shape complementarity statistic and the atom packing are both attempts to measure the importance of the van der Waals forces in an interface one might expect the values to be correlated (LoConte et al., 1999). Neither the average atom volume for completely buried atoms nor for partially buried atoms nor the ratios for these values over the average atom volumes of the protein cores show any correlation with the shape complementarity statistic.

The packing of atoms at interfaces generates histograms with distributions that more closely resemble bimodal distributions than any other distributions in this study (Figure 7). This bimodality is echoed in the complexes between antibodies and antigens, between MHC class I, II molecules, homologs, Fc molecules and ligands, the complexes between proteases and protease inhibitors and the complexes in the G-proteins, cell-cycle, and signal transduction categories. It is intriguing that the buried interface atoms of protease inhibitors and the antigens occupy larger volumes relative to their cores than the proteases and the antibodies. The volumes occupied by atoms in the protein cores
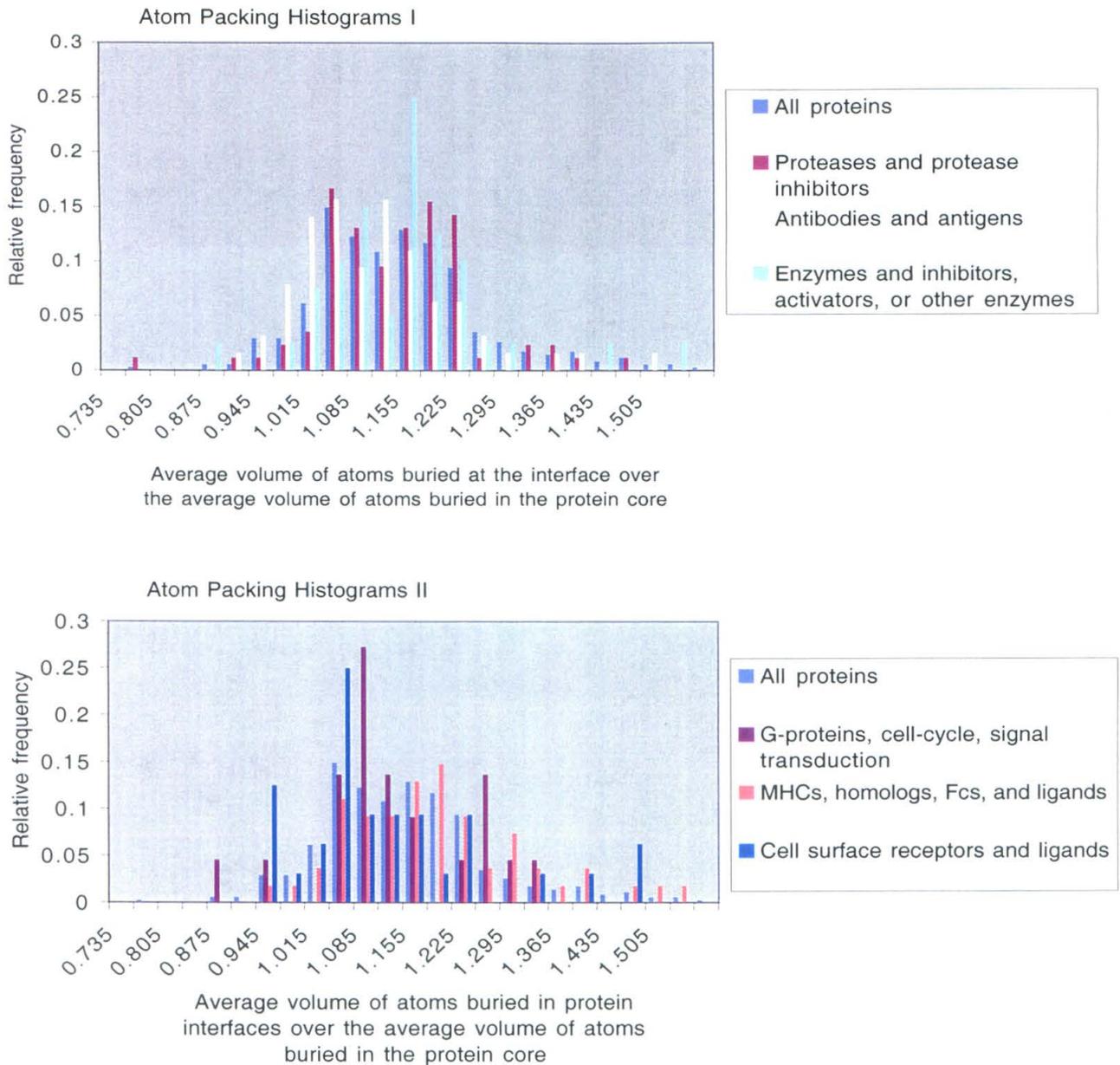
Figure 7: The atom packing histograms are consistently bimodal both as a whole and form the individual classes. This bimodality may indicate that one side of a non- obligate protein-protein interface is usually better packed than the other. This would mean that one side is more stabilized by van der Waals forces than the other.

are more tightly conserved than the other values. These data suggest that the catalytic sites of proteases and the antigen binding sites of antibodies are better packed than the antigens and the protease inhibitors.

Hydrogen Bonds

As the shape complementarity and atom volume addressed the extent of the van der Waals contacts across the interface, we now examine the extent of the electrostatic interactions by assaying the interfaces for the number of hydrogen bonds they contain. To find the hydrogen bonds across an interface, we use Skarzynski and Leslie's algorithm, available as "contact," in the CCP4 suite (CCP4, 1994). We set the program to find all the inter-surface hydrogen bonds between non-carbon atoms. We constrain the bonds found by limiting them in length to greater than 2.3 Å and less than 3.3 Å or 4.0 Å, and by limiting the angle of O$\cdots$H$\cdots$N hydrogen bonds to greater than 120° C, and the H$\cdots$O$\cdots$C bonds to greater than 90° C. These limits generate two sets of hydrogen bonds the strictly allowed (2.3 Å <length<3.3 Å) and the generously allowed (2.3 Å <length<4.0 Å).

The mean number of strictly and generously allowed hydrogen bonds per complex is 13 (standard deviation 9) and 24 (standard deviation 16), respectively. To compare this result with the values generated previously (Jones and Thornton, 1996), we also derived the number of hydrogen bonds per 100 Å^2. Our values, 1.5 (standard deviation 1.2) strictly allowed and 2.7 (standard deviation 2.1) generously allowed hydrogen bonds per 100 Å^2 are higher than those posted by Jones and Thornton (1996) for nonobligate heterocomplexes, 1.1-1.4. Comparing the distributions of hydrogen bonds per complex and per 100 Å^2 yields similar results. The distributions for the

different categories of interfaces overlap more tightly for this metric than they do for the all the others except the atom burial (Figure 8). The complexes between antibodies and antigens and between MHC class I, II molecules, homologs, and ligands and Fc molecules and ligands have slightly more hydrogen bonds. The complexes between proteases and protease inhibitors and the complexes in the G-proteins, cell-cycle, and signal transduction category have slightly fewer hydrogen bonds. This inversion of the results from the shape complementarity statistic indicates a trade-off between employing van der Waals forces and the more electrostatic hydrogen bond.

Atom Type

To consider the chemical *Nature* of the atoms at an interface surface independent of the interactions they form with the atoms on the opposing surface, we examine the percent of the interface surface area attributable to the different kinds of atoms of the protein at the interface. We adopt the convention of dividing the atoms of the polypeptide portions of proteins into nonpolar, polar, and charged categories and the definition of these atom types from previous work (Janin et al., 1988). This atom-by-atom analysis classed the protein atoms observed in crystal structures into three groups: all carbon atoms were considered nonpolar, all oxygen, nitrogen, and sulfur atoms were considered polar, except the carboxylate oxygens, and the amino and guanadinium nitrogen atoms that were considered charged. To provide context for understanding the histograms for these data, we introduce a plot (Figure 9) demonstrating how dividing the atoms of a protein into these categories describes five parts of the proteins in this study. We measure the percent of the total atoms, percent of the core atoms, percent of the total
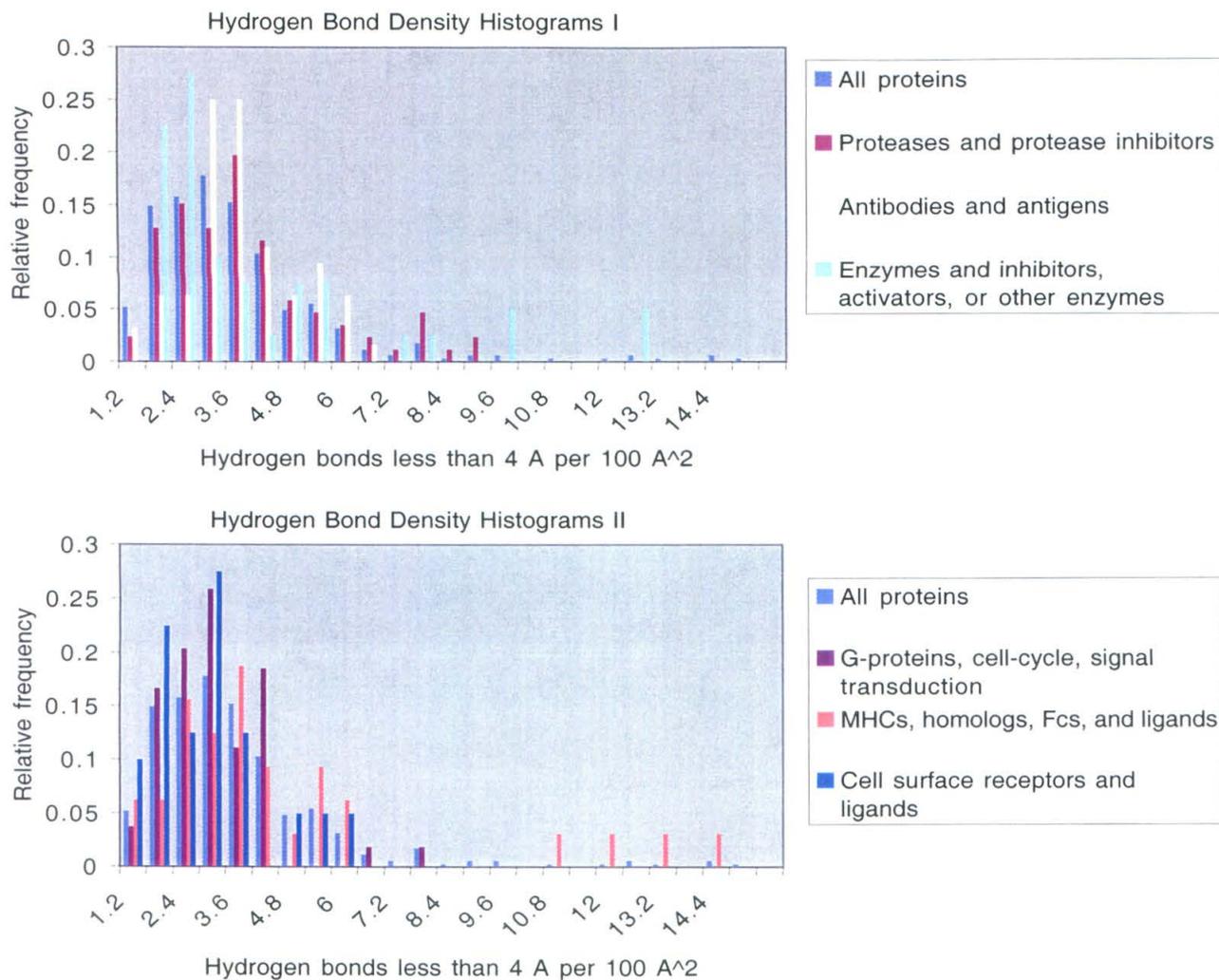
Figure 8: The hydrogen bond density histograms reveal distributions similar to those seen for atom burial. The various biological categories in our study use characteristic numbers of hydrogen bonds.
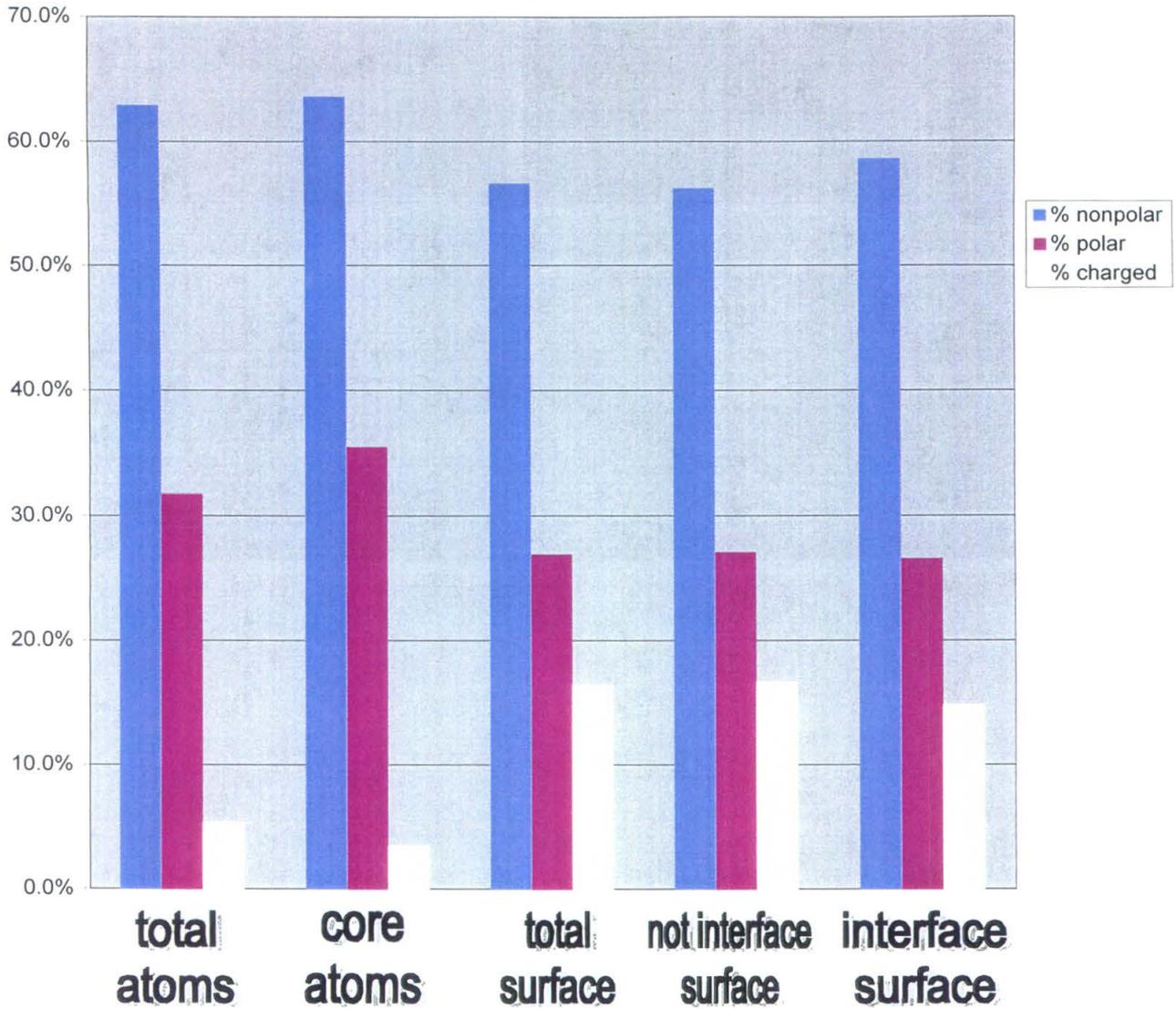
Figure 9: The atom type histograms reveal consistency in the relative amounts of different atom types or surface area attributed to atom types for the different parts of the protein. One significant difference is the lower proportion of charged atoms in the cores of proteins.

surface area, percent of the surface area not involved in known interfaces and the percent of the interface surface area attributable to nonpolar, polar, and charged atoms. It may be observed that the number of nonpolar atoms dominates proteins generally at around 60%. The exclusion of charged atoms from the core of proteins has been noted extensively in the literature (Janin et al., 1988) and is shown in the second trio of columns. We note that for the surface where all atoms are being weighted by their degree of exposure, the interface surface has slightly less charged surface and slightly more nonpolar surface than the surface does generally.

The interfaces of our study are 58.6% (standard deviation 7.6%) nonpolar, 26.6% (standard deviation 8.4%) polar, and 14.8% (standard deviation 8.8%) charged. These values agree well with the 56%, 29%, and 15%, observed in previous study that used 75 complexes (Lo Conte et al., 1999). Similarly, the changes in the proportions of the atom types from category to category of interface agree well with their results. What is apparent in the histograms here that is more difficult to glean from their presentation is that the categories segregate well by the percent of their surfaces that is nonpolar, less well by the polar atom type and not well by the charged atom (Figure 10). The antibodies and antigens interface surfaces have a lower percentage of nonpolar atoms than the mean. The proteases and protease inhibitor, and G-protein, cell-cycle, signal transduction interfaces have a higher nonpolar percentage than the mean. The cell-surface receptors and ligands and the combined categories of MHC class I, II molecules, their homologues, Fcs, and ligands are spread broadly about the mean. These data suggest that along with the circularity and the shape complementarity, the mean of the percent of the interface
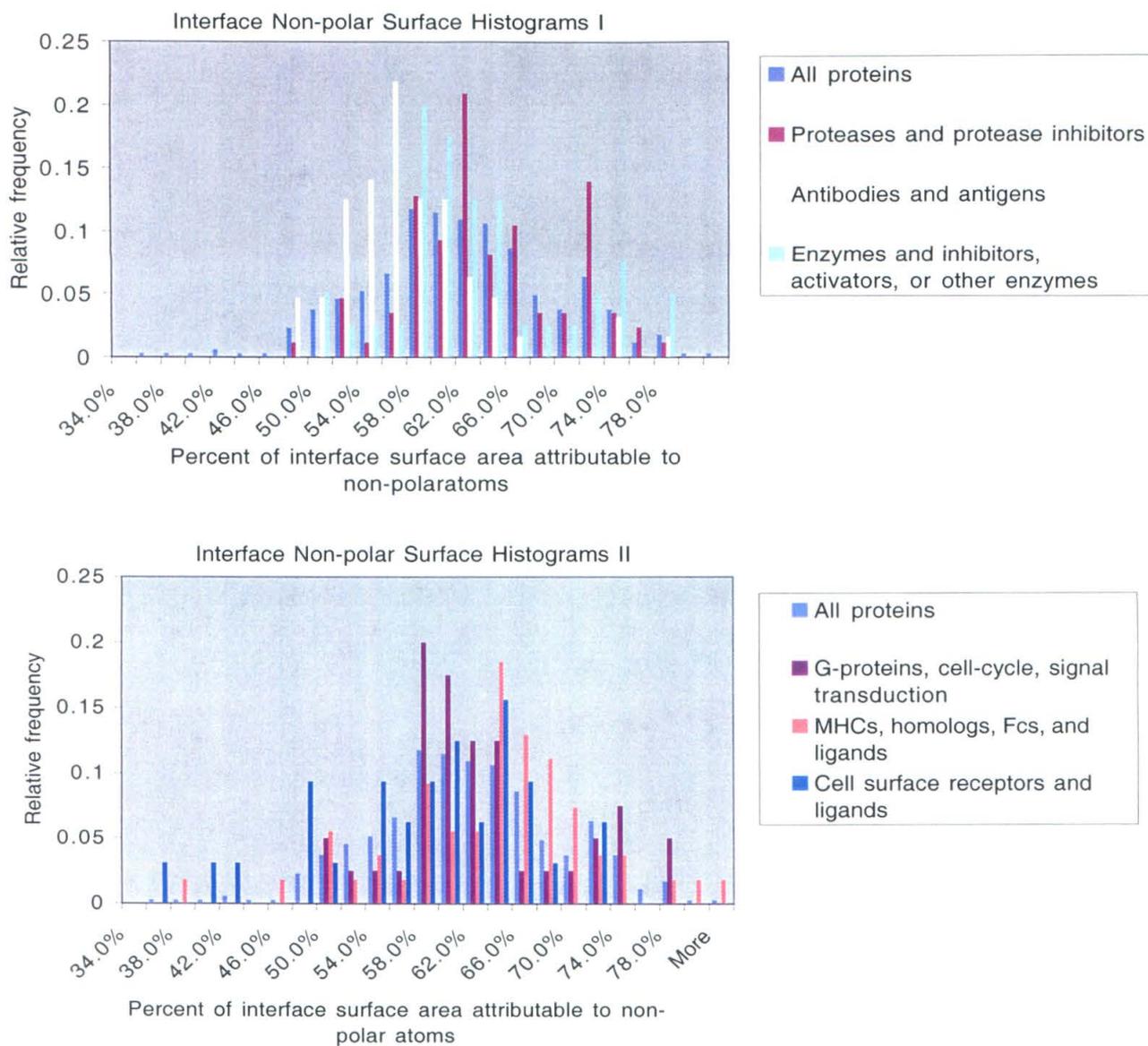
Figure 10a: These atom type histograms show that nonpolar atoms are present in levels more consistent with the various categories of our study than the charged or polar atoms.
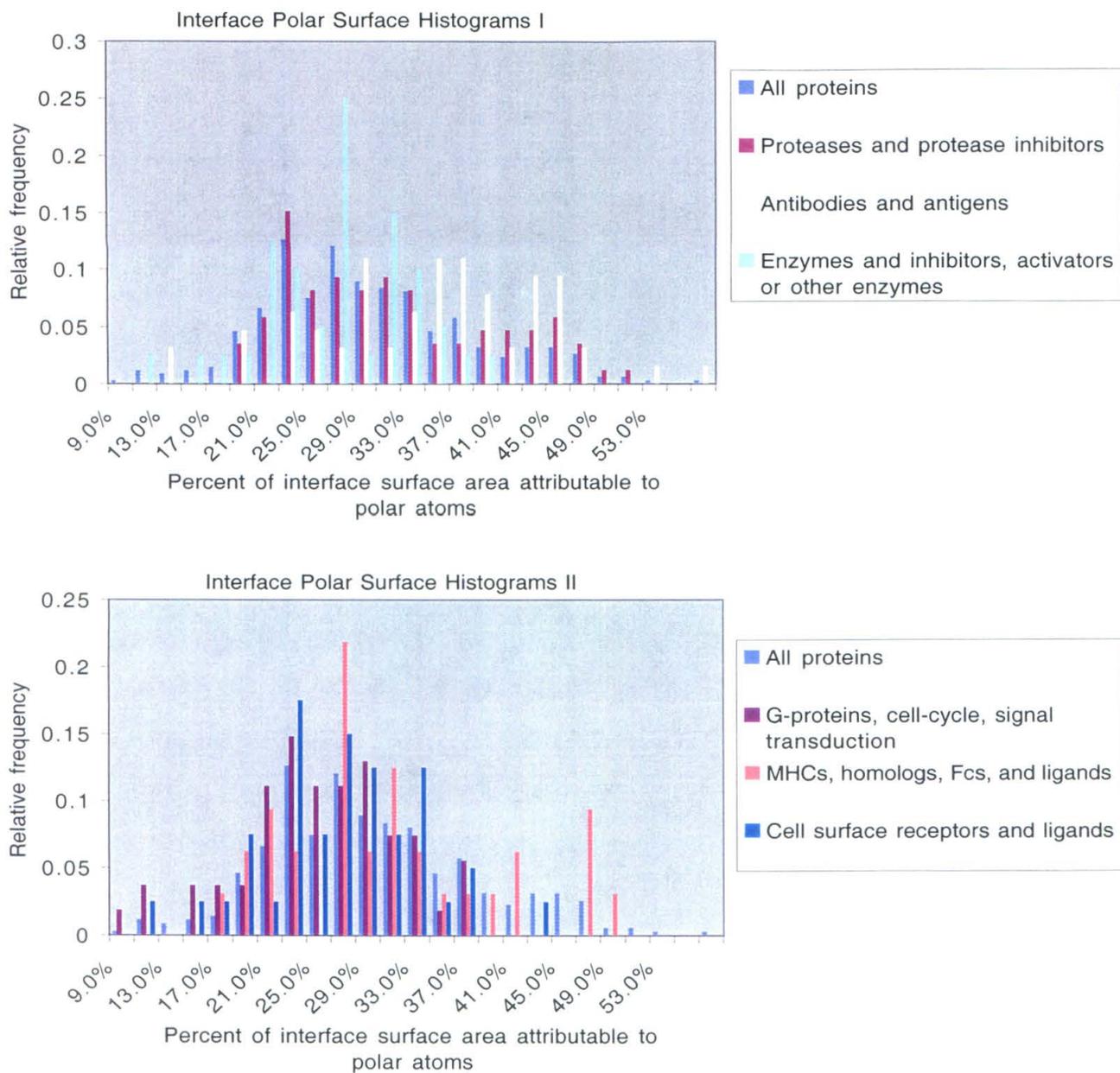
Figure 10b: These atom type histograms show that polar atoms are present in levels less consistent with the various categories of our study than the nonpolar atoms.

Interface Charged Surface Histograms I

Legend: All proteins; Proteases and protease inhibitors; Antibodies and antigens; Enzymes and inhibitors, activators, or other complexes

Percent of interface surface area attributable to charged atoms

Interface Charged Surface Histograms II

Legend: All proteins; G-proteins, cell-cycle, signal transduction; MHCs, homologs, Fcs, and ligands; Cell surface receptors and ligands

Percent of interface surface area attributable to charged atoms

Figure 10c: These atom type histograms show that charged atoms are present in levels more consistent with the various categories of our study than the nonpolar atoms.

surface attributable to nonpolar atoms is probably stable and unlikely to be shifted by additional data.

It is worth noting that decomposing the protein-level metric, the interface surface area, into the amount of interface surface area attributed to nonpolar, polar, and charged atoms we are, with the percent nonpolar surface, explicitly isolating the proportion of the protein-level metric that should be responsible for the hydrophobic effect. It is possible that the more consistent use of nonpolar surface area with the protein categories indicates that this percentage is more necessary to conserve than the polar and charged surface areas which may be more interchangeable.

Amino Acid Composition

The final metric we probed interface surfaces with is the percent of the surface area that may be attributable to different amino acids. Differences in the participation of residues at protein-protein interfaces may be used by protein biochemists to rapidly probe by mutagenesis the surface of a protein to find the protein interfaces. It is known that certain residues (tryptophan, tyrosine, arginine), when mutated to alanine, have on average, a larger effect on the affinity of an interaction than other residues (Bogan and Thorn, 1998). There could be several reasons for this, most obvious is that these are all large residues and the mutation to alanine is a sterically drastic change. The other reasons are subtler. It has been known for a long time that tryptophan and tyrosine, are uncommon (1.3% and 3.2% of all residues) in proteins generally (Creighton, 1984 p.4). It was shown more recently that tryptophan, tyrosine, methionine, cysteine contribute the total surface area of a protein in percentages that are very close to those of their frequency in proteins generally (LoConte et al., 1999). LoConte further showed that

these same four residues contribute much larger percentages to the interfaces of protein-protein interactions and that proline, lysine and glutamate contribute much less surface area to interfaces than they do the surface generally. Our review extends these same results over this larger database.

We present two different charts. One presents the percent contribution a given amino acid makes to the total number of atoms, the number of core atoms, the total surface, the surface not involved in interfaces, and the surface involved in interfaces (Figure 11). From these charts, it may be seen that alanine is the only amino acid that contributes similar percentages to all the portions of the protein. It may also be noted that all those amino acids except arginine that are present at higher percentages on the surface generally are present at lower percentages at interfaces. It may also be observed that the residues that occupy a higher percent of core atoms than they do all atoms also consume a higher percentage of the interface surface area than they do the surface generally. In other words, core residues are also protein-protein interface surface residues, though not so dominantly. Finally, we present a plot that shows the ratio of the percentage of interface surface area occupied by an amino acid over the percentage of the surface area not involved in interface (Figure 12). This plot shows that for the finding the surface of a protein that interacts with another protein by mutagenesis, one has the greatest likelihood of success by scanning the tryptophans, tyrosines, methionines, phenylalanines, cysteines, isoleucines and histidines in that order. The odds are further enhanced if one knows whether the proteins are proteases or immunologically related. The former proteins employ cysteines and methionines at their interfaces on average while the latter do not.

Figure 11: These amino acid histograms reveal how amino acids that are present at higher levels in the cores of proteins are also present at higher levels in nonobligate protein-protein interfaces.
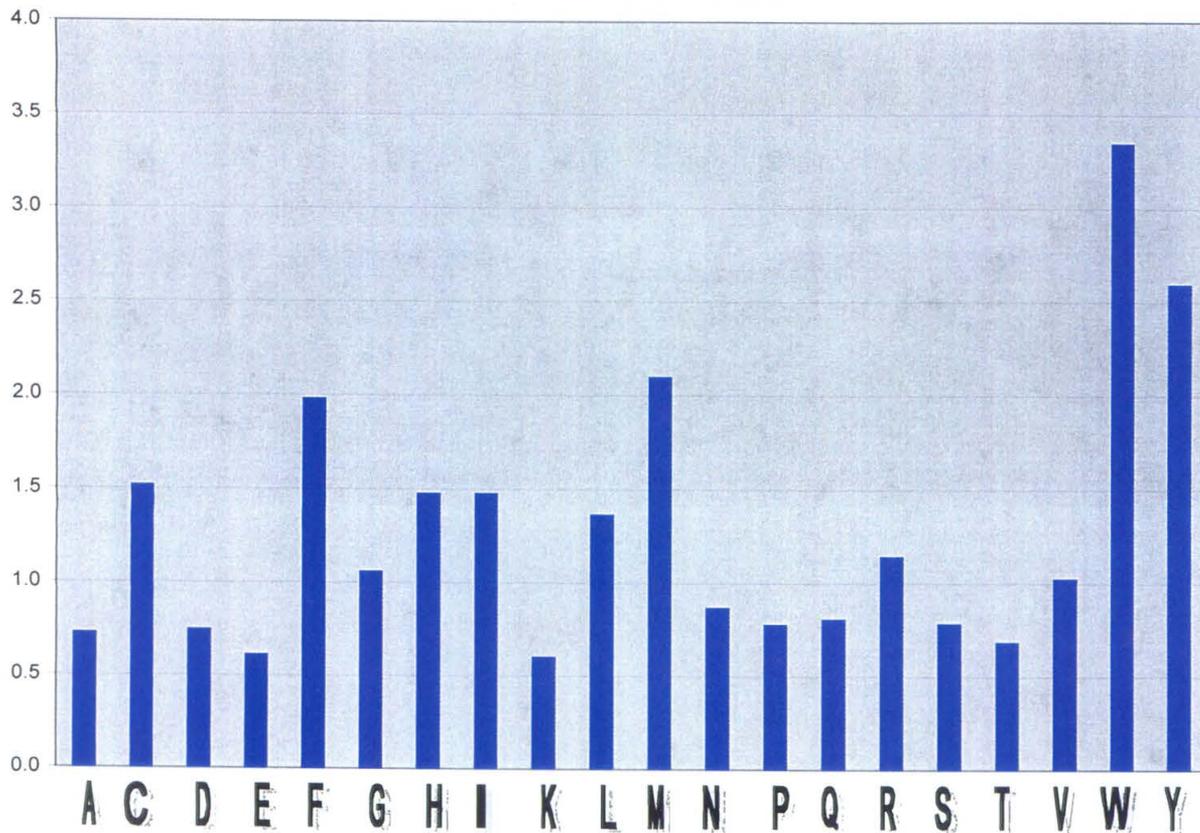
Figure 12: This plot reveals how certain amino acids contribute more surface area to nonobligate protein-protein interfaces than they do to the rest of the surface of these proteins.

Summary: atom level analyses

We have looked at the interfaces of proteins with metrics that assay proteins on the atom-size scale and found, as we did for the protein-size scale, metrics which will serve a variety of purposes. The atom accessibility, like the planarity allows us to distinguish which aspects of a surface are more likely to be involved in protein-protein interfaces. The amount of surface area buried delineates more than any other metric the biologically clustered categories used in this study. The atom-burial helps us describe structural distinctions between the categories of complexes. As the size of the interface surface area gives us a gross assessment of the extent of the hydrophobic effect, the percent of the interface surface area attributable to nonpolar surface may be a more precise measure of the hydrophobic effect. In contrast, the shape complementarity, the atom packing, and the hydrogen bond density allow us to address the extent of the van der Waals and electrostatic contacts between the surfaces. The decomposition of the interface atoms into the chemical categories generated broad distributions for most of the categories, the percent of interface surface area that can be attributed to nonpolar, polar and charged are good descriptors of interfaces. The final metric, the percent of the interface surface area attributable to given amino acids, provides both descriptive and predictive power. Areas of surface with high percentages of tryptophan are more likely to be interface surfaces than those surfaces with high percentages of lysine. Interface surfaces with cysteine and methionine are distinct from most other interfaces.

Correlation analysis

Principal component analysis of the database

We quantitatively described aspects of protein-protein interfaces in a variety of ways and we would like to know if any of these aspects are correlated with any other. To determine to what degree the analyses in this study generated highly correlated measurements of a protein interface, we conducted a principal component analysis of the various analyses over all the samples. In this way, we hope to be able to distinguish the different types of interfaces in a robust mathematical manner. Each metric (interface surface area, circularity, etc.) of each interface (Fab D1.3, lysozyme, etc.) is a dimension. The dimensions of the interfaces are may be arranged horizontally and the dimensions of the metrics may be arranged vertically to construct a sample matrix. Each interface is a row vector and each metric is a column vector of this sample matrix. To conduct a principal component analysis the actual sample values in the sample matrix must be normalized by subtracting the mean observation for that variable from the sample and dividing the difference by the standard deviation. This creates a normalized sample matrix (N (r X c), where r and c are the number of rows and columns). These normalized values have a mean of 0 and a standard deviation of 1. A principal component analysis may be used to minimize the number of column vectors necessary to describe a sample by converting the normalized sample values into values that have no covariance, *i.e.* that are normal to each other. Mathematically this is accomplished by taking the covariance matrix of N. The column vectors of the covariance matrix will have different magnitude and describe different amounts of the variance. To normalize the magnitudes of the vectors and extract the differences in magnitude a singular value decomposition of the

covariance matrix is found. The results of the singular value decomposition of the covariance matrix of the normalized sample matrix used in a PCA are a c X c matrix whose orthonormal column vectors are the eigenvectors of the normalized sample matrix. The proportion of the total variance described by each eigenvector is found in the c X 1 matrix of eigenvalues of the covariance matrix. In this study we used Matlab student release 5.3.0 to find the covariance matrix of the normalized sample matrix, and to do the singular value decomposition of the covariance matrix. We used a graphical program, authored by Ben Bornstein of the Machine Learning Group at JPL, to visualize the results. The three dimensional coordinates in these plots are the product of first three eigenvectors times the normalized sample matrix.

We conducted two principal component analyses over all the interface samples considering the interface surface area, circularity, planarity, accessibility, shape complementarity, Voronoi volume of buried and partially buried atoms, hydrogen bond number, (4 Å distance cutoff), and the percent surface area associated with the nonpolar, polar, and charged atom types. One analysis included the nitrogenase (1n2c, Schindelin et al, 1997) and carbon monoxide (1ffu, Haenzelmann et al., 2000) complexes, these were found to cluster by themselves distant from the other interfaces largely because of their size. The other analysis reported here omitted the interfaces from these two structures. This analysis was subjected to a cross-validation analysis, where the principal component analysis was repeated with one of the variables deleted each time. The first three eigenvectors generated in the first principal component analysis describe collectively 51.4% of the total variance contained in the eleven original measurements (Figure 13). Most of the variance described by the first eigenvector (this vector describes 23.0% of the

total variance) comes from atom accessibility, the percent of charged interface surface, and anticorrelated contributions from the percent polar interface surface. The second eigenvector describes 15.8% of the total variance. The largest contributions of variance from the original parameters to the second eigenvector percent neutral surface area, the Voronoi volumes of atoms buried at the interface and anticorrelated contributions from the percent charged surface area. The third eigenvector accounts for 12.6% of the total variance. It has large contributions from the percent polar surface area and planarity that are anticorrelated to the shape complementarity and the circularity. The forth eigenvector accounts for 11.2% of the total variance and is dominated by the hydrogen bond measurement which is anticorrelated with the shape complementarity.

This principal component analysis is consistent with the individual analyses of the metrics above and is informative about the correlations between the various analyses. The interfaces vary most in terms of the atom accessibility and the percent of the buried surface area attributed to neutral, polar, and charged groups. The anticorrelation between the number of hydrogen bonds and the shape complementarity seen here was noted above. Similar to the energetic trade-off implicated in that relationship, the correlation between the percent nonpolar interface surface and the Voronoi volumes of atoms buried at the interface suggests an exchange between the use of the hydrophobic effect and the use of van der Waal's forces in forming a nonobligate interface.
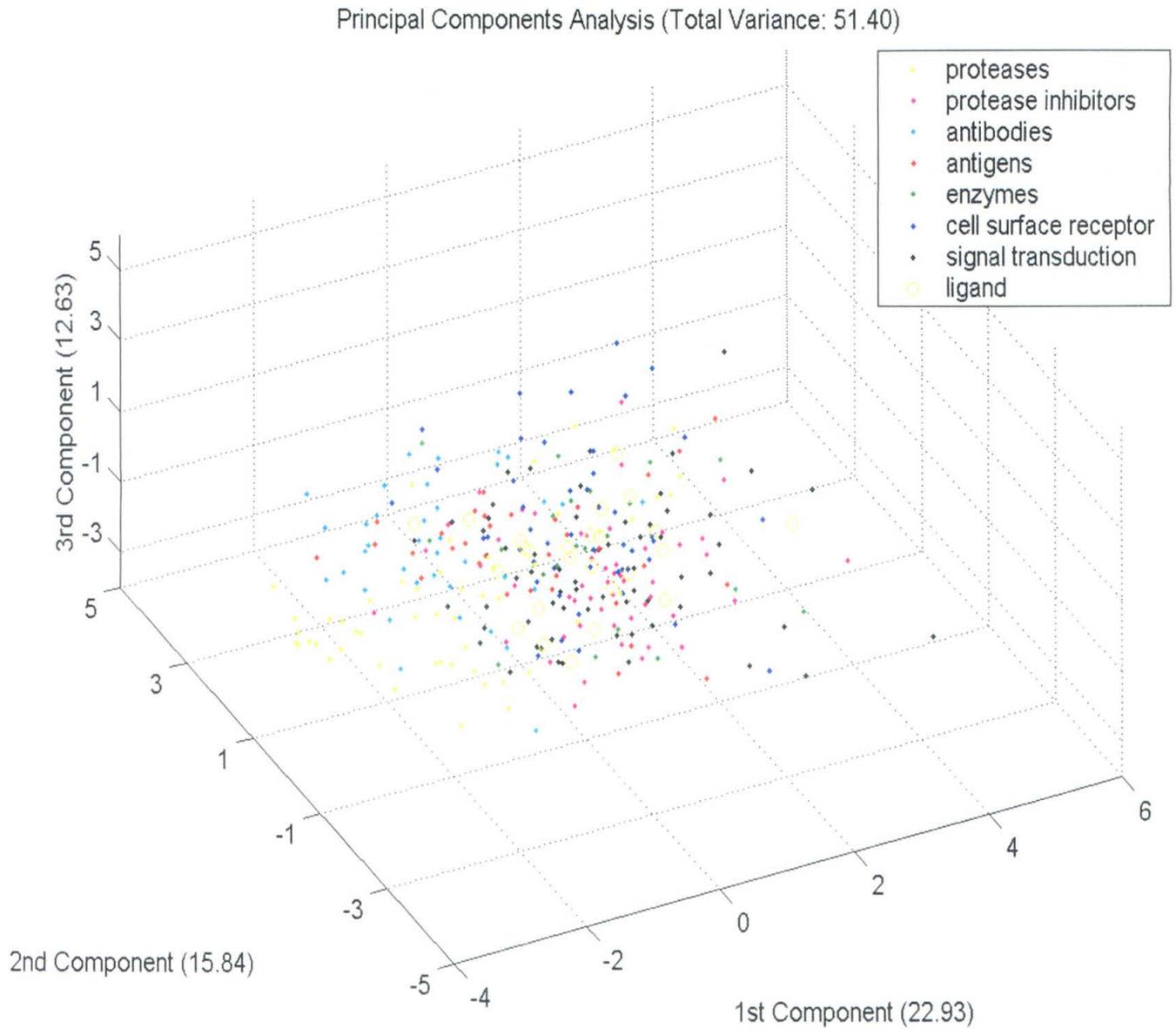
Figure 13: This principal component analysis reveals that for this diverse set of metrics

the proteins in our study to not form separate clusters.

Figure 14: This principal component analysis reveals that for the percent of surface contributed by the different amino acids the proteins in our study do not form separate clusters, also.

A principal component analysis that measured the covariance between the contributions to the interface surface area attributable to different amino acids was also conducted over the complete database and the database without the interfaces from the nitrogenase and carbon monoxidase (Figure 14). For the amino acid percentages, these databases generated identical results. The first three eigenvectors describe 28.2% of the total variance captured in these twenty percentages. The first eigenvector (12.1% of the total variance) has sizable, correlated contributions from tyrosine, tryptophan, asparagine, and serine. The second eigenvector (8.5% of the total variance) contains large correlated contributions from methionine, histidine, and phenylalanine. The third eigenvector (7.5% of the total variance) has correlated contributions from cysteine and arginine and anticorrelated contributions from leucine and valine.

Visualizing the principal component analyses revealed that the categories used in the analyses above might be improved. To view the clustering, there are three-dimensional scatter plots with axes of the three eigenvectors with the largest eigenvalues, that is the three new dimensions that describe the greatest amount of variance. Visualizing the principal component analyses revealed that a change in the category scheme would be informative. The color-coding of the independent analyses is based on the categories outlined at the beginning of the paper. These categories were chosen in part because they were consistent with previous analyses (Jones and Thornton, 1996; LoConte et al, 1999). The new category scheme, shown in the color-coding of the principal component analyses is based on categories that may consist of only one side of an interface. In this new scheme, the proteases are yellow, the protease inhibitors are magenta, antibodies are yellow, antigens are red. The other categories are compressed

into enzymes (green), cell-surface receptors (dark blue), ligands (empty yellow squares), and signal-transduction molecules (black).

Even with the modified category scheme, examining the plots of the principal component analyses reveals that the categories do not form completely separate clusters. These plots show that the categories of interfaces overlap even when the descriptors are orthonormal to each other, even when the axes of the plot maximize the variance between the interfaces they to not form separate clusters. This means that in most cases the different categories of proteins complexes form interfaces in largely the same way. There are however two interesting exceptions. It may be seen from these plots is that the interfaces of proteases (yellow) and antibodies (light blue) generate a scatter that has a center removed from that of the general scatter. In other words, based on the measurements in this study the interfaces surfaces of proteases and antibodies are special cases of interfaces and distinctly not representative of the average protein interfaces. That they are distinct is intriguing because they were the first structures to be determined (Janin and Chothia, 1990) and they have shaped our understanding of what protein-protein interfaces are.

Conclusions

Examining a large data set with this broad set of commonly used metrics largely confirms and extends previous attempts to examine protein-protein interfaces. The protein-level analyses describe the shape and size of the interface and in the interface surface area provide a crude method of estimating the role of the hydrophobic effect in the association of the proteins. As has been noted earlier (Jones and Thornton, 1997b), the protein-level metric, planarity, and the atom-level metric, atom accessibility, are

aspects of interface surfaces that are different on average from the protein surface generally. The atom-level analyses are more descriptive of the chemical forces involved in forming interfaces. The shape complementarity and atom packing are ways of estimating the van der Waals forces. The hydrogen bond density as well as the percent of the interface surface attributable to charged atoms might be descriptive of the electrostatic component of the interaction. The percent of the interface surface attributable to nonpolar atoms is probably a more precise method of describing the role of the hydrophobic effect than the size of the interface surface area. The amino-acid composition of protein-protein interfaces reveals that the amino acids more likely to be present in protein cores than in a protein in general are also more likely to be present at interfaces than at the surface in general. The surface of interfaces attributed to certain amino acids, tryptophan, tyrosine, and methionine, is a greater percentage of interface surface area than these two amino acids contribute to surface area generally. The opposite is true for lysine, proline, and glutamate. By using the trends examined here in amino acid composition, planarity, and atom accessibility, a protein biochemist may more rapidly identify the surface of a protein involved in interacting with another protein.

The principal component analyses of the interface metrics are original to this work, they corroborate the previous analyses and illuminate several questions that may be worth further consideration. The principal component analysis conducted on the interface surface area, circularity, planarity, accessibility, shape complementarity, packing of buried and partially buried atoms, hydrogen bond number, (4 Å distance cutoff), and the percent surface area associated with the nonpolar, polar, and charged atom types, explained 51.4% of the total variance with three variables. This analysis

confirmed the ability of the atom accessibility and the percent of interface surface area attributable to the different atom types to describe the variance seen in protein interfaces. This analysis also confirmed an anticorrelation noted between the surface complementarity and the number of hydrogen bonds and discovered a correlation between the amount of nonpolar interface surface and the Voronoi volume of atoms buried at the interface surface that is indicative of a similar energetic exchange. The principal component analysis of the percent interface surface attributable to different amino acids showed that tryptophan, tyrosine, serine, and asparagine are responsible for a large amount of the variance in protein interfaces.

Visualizing the principal component analyses showed three things. The principal component analyses showed how the categories of interfaces could be informative if the categories can contain only the related halves of nonobligate complexes. Both principal component analyses revealed that the interfaces do not cluster into completely separate groups. However, these analyses also reveal the proteases and the antibodies are more removed from the general cluster of interfaces. As such they represent somewhat exceptional instead of typical nonobligate protein-protein complexes.

References

Ay, J., Keitel, T., Kuttner, G., Wessner, H., Scholz, C., Hahn, M., Hohne, W. (2000) 1DZB, Crystal structure of a phage library-derived single-chain Fv fragment complexed with turkey egg-white lysozyme at 2.0 A resolution. *J. Mol. Biol.* 301 pp. 239-246.

Banner, D. W., D'Arcy, A., Janes, W., Gentz, R., Schoenfeld, H. J., Broger, C., Loetscher, H., Lesslauer, W. (1993) 1TNR, Crystal structure of the soluble human 55 kd TNF receptor-human TNF beta complex: implications for TNF receptor activation. *Cell* 73, 431-445.

Ban, N., Escobar, C., Garcia, R., Hasel, K., Day, J., Greenwood, A., McPherson, A. (1994) 1IAI, Crystal structure of an idiotype-anti-idiotype Fab complex. *Proc Natl Acad Sci USA* 91, 1604-1608.

Banner, D. W., D'Arcy, A., Chene, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y., Kirchhofer, D. (1996) 1DAN, The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature* 380, 41-46.

Bennett, M. J., Lebron, J. A., Bjorkman, P. J. (2000) 1DE4, Crystal structure of the hereditary haemochromatosis protein hfe complexed with transferrin receptor. *Nature* 403, 46-53.

Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F. P., Mariuzza, R. A., Poljak, R. J. (1994) 1VFB, Bound water molecules and conformational stabilization help mediate an antigen-antibody association. *Proc Natl Acad Sci USA* 91, 1089-1093.

Bode, W., Wei, A. Z., Huber, R., Meyer, E., Travis, J., Neumann, S. (1986) 1PPF, X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor. *EMBO J* 5, 2453-2458.

1CSE, Bode, W., Papamokos, E., Musil, D. (1987) The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech Hirudo medicinalis. Structural analysis, subtilisin structure and interface geometry. *Eur. J. Biochem.* 166, 673-692.

Bode, W., Greyling, H. J., Huber, R., Otlewski, J., Wilusz, T. (1989) 1PPE, The refined 2.0 A X-ray crystal structure of the complex formed between bovine beta-trypsin and CMTI-I, a trypsin inhibitor from squash seeds (Cucurbita maxima). Topological similarity of the squash seed inhibitors with the carboxypeptidase A inhibitor from potatoes. *FEBS Lett* 242, 285-292.

Bogan, A. A. and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* 280, 1-9.

Bolognesi, M., Gatti, G., Menagatti, E., Guarneri, M., Marquart, M., Papamokos, E., Huber, R. (1982) 1TGS, Three-dimensional structure of the complex between pancreatic secretory trypsin inhibitor (Kazal type) and trypsinogen at 1.8 A resolution. Structure solution, crystallographic refinement and preliminary structural interpretation. *J. Mol. Biol.* 162, 839-868.

Bompard-Gilles, C., Rousseau, P., Rouge, P., Payan, F. (1996) 1DHK, Substrate mimicry in the active center of a mammalian alpha-amylase: structural analysis of an enzyme-inhibitor complex. *Structure* 4, 1441-1452.

Bossart-Whitaker, P., Chang, C. Y., Novotny, J., Benjamin, D. C., Sheriff, S. (1995) 1NSN, The crystal structure of the antibody N10-staphylococcal nuclease complex at 2.9 A resolution. *J. Mol. Biol.* 253, 559-575.

Boyington, J. C., Motyka, S. A., Schuck, P., Brooks, A. G., Sun, P. D. (2000) 1EFX, Structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. *Nature* 405, 537-543.

Braden, B. C., Souchon, H., Eisele, J. L., Bentley, G. A., Bhat, T. N., Navaza, J., Poljak, R. J. (1994) 1MLC, Three-dimensional structures of the free and the antigen-complexed Fab from monoclonal anti-lysozyme antibody D44.1. *J. Mol. Biol.* 243, 767-781.

Braden, B. C., Fields, B. A., Ysern, X., Dall'Acqua, W., Goldbaum, F. A., Poljak, R. J., Mariuzza, R. A. (1996) 1DVF, Crystal structure of an Fv-Fv idiotope-anti-idiotope complex at 1.9 A resolution. *J. Mol. Biol.* 264, 137-151.

Buckle, A. M., Schreiber, G., Fersht, A. R. (1994) 1BRS, Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-A resolution. *Biochemistry* 33, 8878-8889.

Burgering, M. J., Orbons, L. P., van der Doelen, A., Mulders, J., Theunissen, H. J., Grootenhuis, P. D., Bode, W., Huber, R., Stubbs, M. T. (1997) 1TFX, The second Kunitz domain of human tissue factor pathway inhibitor: cloning, structure determination and interaction with factor Xa. *J. Mol. Biol.* 269, 395-407.

Chacko, S., Silverton, S., Hibbits, K., Xavier, A., Willson, R. and Davies, D. (1995) 1BQL, Structure of an Anti-Hel Fab Fragment Complexed with Bobwhite Quail Lysozyme, To be Published.

Celikel, R., Ruggeri, Z. M., Varughese, K. I. (2000) 1FNS, Modulation of Von Willebrand factor conformation and adhesive function by an internalized water molecule. *Nat. Struct. Biol.* 7, 189-194.

Chen, L., Glover, J. N., Hogan, P. G., Rao, A., Harrison, S. C. (1998) 1A02, Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature* 392, 42-48.

Chen, Z., Bode, W. (1983) 2KAI, Refined 2.5 Angstroms X-ray crystal structure of the complex formed by porcine kallikrein a and the bovine pancreatic trypsin inhibitor. crystallization, patterson search, structure determination, refinement, structure and comparison with its components and with the bovine trypsin-pancreatic trypsin inhibitor complex. *J. Mol. Biol.*, 164, 283-311.

Chitarra, V., Alzari, P. M., Bentley, G. A., Bhat, T. N., Eisele, J. L., Houdusse, A., Lescar, J., Souchon, H., Poljak, R. J. (1993) 1JHL, Three-dimensional structure of a heteroclitic antigen-antibody cross-reaction complex. *Proc Natl Acad Sci USA* 90, 7711-7715.

Clackson T. and J. A. Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* 267, 383-386.

Cohen, G. H., Sheriff, S., Davies, D. R. (1996) 3HFL, Refined structure of the monoclonal antibody HyHEL-5 with its antigen hen egg-white lysozyme. *Acta Cryst. D* 52, 315-326.

Collaborative Computational Project, Number 4. (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst. D* 50, 760-763.

Corper, A. L., Sohi, M. K., Bonagura, V. R., Steinitz, M., Jefferis, R., Feinstein, A., Beale, D., Taussig, M. J., Sutton, B. J. (1997) 1ADQ, Structure of human IgM

rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction. *Nat. Struct. Biol.* 4, 374-381.

Creighton, T. E. (1993) Proteins: structures and molecular properties. W. H. Freeman and Co, New York, NY.

Cunningham, B. C. and Wells, J. A. (1993) Comparison of a structural and functional epitope. *J. Mol. Biol.* 234, 554-563.

Dauter, Z., Betzel, C., Genov, N., Pipon, N., Wilson, K. S. (1991) 1MEE, Complex between the subtilisin from a mesophilic bacterium and the leech inhibitor eglin-C. *Acta Cryst B* 0, 707-730.

Deisenhofer, J. (1981) 1FC2, Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution. *Biochemistry* 20, 2361-2370.

Deisenhofer, J. (1981) 1FC1, Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution. *Biochemistry* 20, 2361-2370.

Derrick, J. P., Wigley, D. B. (1994) 1IGC, The third IgG-binding domain from streptococcal protein G. An analysis by X-ray crystallography of the structure alone and in a complex with Fab. *J. Mol. Biol.* 243, 906-918.

Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., Muyldermans, S., Wyns, L. (1996) 1MEL, Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat. Struct. Biol.* 3, 803-811.

Elkins, P. A., Christinger, H. W., Sandowski, Y., Sakal, E., Gertler, A., De Vos, A. M., Kossiakoff, A. A. (2000) 1F6F, Ternary complex between placental lactogen and the extracellular domain of the prolactin receptor. *Nat. Struct. Biol.* 7, 808-815.

Engh, R. A., Girod, A., Kinzel, V., Huber, R., Bossemeyer, D. (1996) 1YDR, Crystal structures of catalytic subunit of cAMP-dependent protein kinase in complex with isoquinolinesulfonyl protein kinase inhibitors H7, H8, and H89. Structural implications for selectivity. *J. Biol. Chem.* 271, 26157-26164.

Fields, B. A., Malchiodi, E. L., Li, H., Ysern, X., Stauffacher, C. V., Schlievert, P. M., Karjalainen, K., Mariuzza, R. A. (1996) 1JCK, Crystal structure of a T-cell receptor beta-chain complexed with a superantigen. *Nature* 384, 188-192.

Fleury, D., Wharton, S. A., Skehel, J. J., Knossow, M., Bizebard, T. (1998) 1VIR, Antigen distortion allows influenza virus to escape neutralization. *Nat. Struct. Biol.* 5, 119-123.

Fleury, D., Daniels, R. S., Skehel, J. J., Knossow, M., Bizebard, T. (2000) 1EO8, Structural evidence for recognition of a single epitope by two distinct antibodies. *Proteins: Struct., Funct., Genet.* 40, 572-578.

Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P., Bolognesi, M. (1992) 1ACB, Crystal and molecular structure of the bovine alpha-chymotrypsin-eglin c complex at 2.0 A resolution. *J. Mol. Biol.* 225, 107-123.

Fuentes-Prior, P., Noeske-Jungblut, C., Donner, P., Schleuning, W. D., Huber, R., Bode, W. (1997) 1AVG, Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug. *Proc Natl Acad Sci USA* 94, 11845-11850.

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., M. Laskowski Jr., James, M. N. (1987) 1CHO, Crystal and molecular structures of the complex of alpha-chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 A resolution. *J. Mol. Biol.* 195, 397-418.

Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., Hill, C. P. (1996) 1AK4, Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* 87, 1285-1294

Gao, G. F., Tormo, J., Gerth, U. C., Wyer, J. R., McMichael, A. J., Stuart, D. I., Bell, J. I., Jones, E. Y., Jakobsen, B. K. (1997) 1AKJ, Crystal structure of the complex between human CD8alpha(alpha) and HLA-A2. *Nature* 387, 630-634.

Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W. E., Wiley, D. C. (1996) 1AO7, Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384, 134-141.

Garcia, K. C., Degano, M., Pease, L. R., Huang, M., Peterson, P. A., Teyton, L., Wilson, I. A. (1998) 2CKB, Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* 279, 1166-1172.

Garman, S. C., Wurzburg, B. A., Tarchevskaya, S. S., Kinet, J. P., Jardetzky, T. S. (2000) 1F6A, Structure of the Fc Fragment of Human Ige Bound to its High-Affinity Receptor FcϵRiα. *Nature* 406, 259-266.

Gerstein , M. (1992) A resolution-sensitive procedure for comparing protein surfaces and its application to the comparison of antigen-combining sites. Acta Cryst. A 48: 271-276.

Gerstein, M., Tsai, J. and M. Levitt, M. (1995) The volume of atoms on the protein surface: Calculated from simulation, using Voroni polyhedra. *J. Mol. Biol.* 249, 955-966.

Gillmor, S. A., Takeuchi, T., Yang, S. Q., Craik, C. S., Fletterick, R. J. (2000) 1EZS, Compromise and accomodation in ecotin, a dimeric macromolecular inhibitor of serine proteases. *J. Mol. Biol.* 299, 993-1003.

Gorina, S., Pavletich, N. P. (1996) 1YCS, Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science* 274, 1001-1005.

Griffith, J. P., Kim, J. L., Kim, E. E., Sintchak, M. D., Thomson, J. A., Fitzgibbon, M. J., Fleming, M. A., Caron, P. R., Hsiao, K., Navia, M. A. (1995) 1TCO, X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. *Cell* 82, 507-522.

Gros, P., Fujinaga, M., Dijkstra, B. W., Kalk, K. H., Hol, W. G. (1989) 1TEC, Crystallographic refinement by incorporation of molecular dynamics: thermostable serine protease thermitase complexed with eglin c. *Acta Cryst B* 45, 488-499, 1989.

Haenzelmann, P., Dobbek, H., Gremer, L., Huber, R., Meyer, O. (2000) 1FFU, The Effect of intracellular molybdenum in Hydrogenophaga pseudoflava on the crystallographic structure of the seleno-molybdo-iron-sulfur flavoenzyme carbon monoxide dehydrogenase. *J. Mol. Biol.* 301, 1223-1233.

Hage, T., Sebald, W., Reinemer, P. (1999) 1IAR, Crystal structure of the interleukin-4 / receptor alpha chain complex reveals a mosaic binding interface. *Cell* 97, 271-281.

Hanson, M. A., Stevens, R. C. (2000) 1F83, Cocrystal structure of synaptobrevin-ii bound to botulinum neurotoxin type b at 2.0 angstrom resolution. *Nat. Struct. Biol.* 7, 687-692.

Harel, M., Kleywegt, G. J., Ravelli, R. B., Silman, I., Sussman, J. L. (1995) 1FSS, Crystal structure of an acetylcholinesterase-fasciculin complex: interaction of a three-fingered toxin from snake venom with its target. *Structure* 3, 1355-1366.

Harrenga, A., Michel, H. (1999) 1QLE, The cytochrome c oxidase from paracoccus denitrificans does not change the metal center ligation upon reduction. *J. Biol. Chem.* 274, 33296-33299.

Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F., Kuriyan, J. (1997) 1DKG, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* 276, 431-435.

Hecht, H. J., Szardenings, M., Collins, J., Schomburg, D. (1991) 1CGJ, Three-dimensional structure of the complexes between bovine chymotrypsinogen A and two recombinant variants of human pancreatic secretory trypsin inhibitor (Kazal-type). *J. Mol. Biol.* 220, 711-722.

Hecht, H. J., Szardenings, M., Collins, J., Schomburg, D. (1991) 1CGI, Three-dimensional structure of the complexes between bovine chymotrypsinogen A and two

recombinant variants of human pancreatic secretory trypsin inhibitor (Kazal-type). *J. Mol. Biol.* 220, 711-722.

Hennecke, J., Carfi, A., Wiley, D. C. (2000) 1FYT, Structure of a covalently stabilized complex of a human αβ-T cell receptor, influenza HA peptide and MHC Class II molecule, Hla-Dr1. *EMBO J* 19, 5611-5624.

Hermoso, J., Pignol, D., Kerfelec, B., Crenon, I., Chapus, C., Fontecilla-Camps, J. C. (1996) 1ETH, Lipase activation by nonionic detergents. The crystal structure of the porcine lipase-colipase-tetraethylene glycol monooctyl ether complex. *J. Biol. Chem.* 271, 18007-18016.

Holmes, M. A., Buss, T. N., Foote, J. (1998) 1BVK, Conformational correction mechanisms aiding antigen recognition by a humanized antibody. *J Exp Med* 187, 479-485.

Hosfield, C. M., Elce, J. S., Davies, P. L., Jia, Z. (1999) 1DF0, Crystal Structure of Calpain Reveals the Structural Basis for Ca2+-Dependent Protease Activity and a Novel Mode of Enzyme Activation. *EMBO J* 18, 6880-6886.

Housset, D., Mazza, G., Gregoire, C., Piras, C., Malissen, B., Fontecilla-Camps, J. C. (1997) 1KB5, The three-dimensional structure of a T-cell antigen receptor V alpha V beta heterodimer reveals a novel arrangement of the V beta domain. *EMBO J* 16, 4205-4216.

Huang, K., Strynadka, N. C., Bernard, V. D., Peanasky, R. J., James, M. N. (1994) 1EAI, The molecular structure of the complex of Ascaris chymotrypsin/elastase inhibitor with porcine elastase. Structure 2, 679-89.

Huang, L., Hofer, F., Martin, G. S., Kim, S. H. (1998) 1LFD, Structural basis for the interaction of Ras with RalGDS. *Nat. Struct. Biol.* 5, 422-426.

Huang, Q., Liu, S., Tang, Y. (1993) 1MCT, Refined 1.6 A resolution crystal structure of the complex formed between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the squash family. Detailed comparison with bovine beta-trypsin and its complex. *J. Mol. Biol.* 229, 1022-1036.

Huntington, J. A., Read, R. J., Carrell, R. W. (2000) 1EZX, Structure of a Serpin-Protease Complex Shows Inhibition by Deformation. *Nature* 407, 923-926.

Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W., Remington, S. J. (1993) 1GLA, Structure of the regulatory complex of Escherichia coli IIIGlc with glycerol kinase. *Science* 259, 673-677.

Huse, M., Chen, Y. G., Massague, J., Kuriyan, J. (1999) 1B6C, Crystal structure of the cytoplasmic domain of the type I TGF beta receptor in complex with FKBP12. *Cell* 96, 425-436.

Hymowitz, S. G., Christinger, H. W., Fuh, G., O'Connell, M. P., Kelley, R. F., Ashkenazi, A., De Vos, A. M. (1999) 1D0G, Triggering cell death: crystal structure of apo2l/trail in complex with death receptor-5. *Mol. Cell* 4, 563-571.

Janin, J. and Chothia, C. (1990) The structure of protein-protein recognition sites. *J. Biol. Chem.* 265, 16027-16030.

Janin, J., Miller, S. and Chothia, C. (1988) Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* 204, 155-164.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., Pavletich, N. P. (1995) 1FIN, Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376, 313-320.

Jeffrey, P. D., Tong, L., Pavletich, N. P. (2000) 1G3N, Structural basis of inhibition of cdk-cycin complexes by ink4 inhibitors, *Genes Dev.* 24, 3115-3120.

Jones, S. and Thornton, J. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* 93, 13-20.

Jones, S. and Thornton, J. (1997a) Analysis of protein-protein interaction sites using surface patches.*J. Mol. Biol.* 272, 121-132.

Jones, S. and Thornton, J. (1997b) Predicition of protein-protein interaction sites using patch analysis. *J. Mol. Biol.* 272, 133-143.

Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., Holmes, K. C. (1990) 1ATN, Atomic structure of the actin:DNase I complex. *Nature* 347, 37-44.

Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., Leberman, R. (1996) 1EFU, The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution. *Nature* 379, 511-518.

Kleywegt, G. J. and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Cryst D* 50, 178-185.

Kleywegt, G. J. (1999) Experimental assessment of differences between related protein crystal structures. *Acta Cryst. D* 55, 1878-1884 1999.

Kern, P. S., Teng, M. K., Smolyar, A., Liu, J. H., Liu, J., Hussey, R. E., Spoerl, R., Chang, H. C., Reinherz, E. L., Wang, J. H. (1998) 1BQH, Structural basis of CD8 coreceptor function revealed by crystallographic analysis of a murine CD8 alpha ectodomain fragment in complex with H-2Kb. *Immunity*, 9, 519-530.

Kirsch, T., Sebald, W., Dreyer, M. K. (2000) 1ES7, Crystal Structure of the Bmp-2-Bria Ectodomain Complex, *Nat. Struct. Biol.* 7, 492-496.

Kobe, B., Deisenhofer, J. (1995) 1DFJ, A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* 374, 183-186.

Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., Hendrickson, W. A. (1998) 1GC1, Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393, 648-659.

Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E., Sigler, P. B.(1996) 1GOT, The 2.0 A crystal structure of a heterotrimeric G protein. *Nature* 379, 311-319.

Lawrence, M. C. and Colman, P. (1993) Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* 234, 946-950.

Lee, C. H., Saksela, K., Mirza, U. A., Chait, B.T., Kuriyan, J. (1996) 1EFN, Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell* 85, 931-942.

Lennon, B. W., , C. H. Williams Jr., Ludwig, M. L. (2000) 1F6M, Twists in Catalysis: Alternating Conformations of Escherichia Coli Thioredoxin Reductase. *Science* 289, 1190-1194.

Lescar, J., Pellegrini, M., Souchon, H., Tello, D., Poljak, R. J., Peterson, N., Greene, M.,

Alzari, P. M. (1995) 1FBI, Crystal structure of a cross-reaction complex between Fab F9.13.7 and guinea fowl lysozyme. *J. Biol. Chem.* 270, 18067-18076.

Lewis, M. and Rees, D. C. (1985) Fractal surfaces of proteins. *Science* 230, 1163-1164.

Li, H., Dunn, J. J., Luft, B. J., Lawson, C. L. (1997) 1OSP, Crystal structure of Lyme disease antigen outer surface protein A complexed with a Fab. *Proc Natl Acad Sci USA* 94, 3584-3589.

Li, Y., Li, H., Smith-Gill, S. J., Mariuzza, R. A. (2000) 1DQJ, Three-dimensional structure of the free and antigen-bound fab from monoclonal antilysozyme antibody Hyhel-63. *Biochemistry* 39, 6296-6309.

Li, Y., Li, H., Dimasi, N., Mccormick, J. K., Martin, R., Schuck, P., Schlievert, P., Mariuzza, R. (2001) 1HQR, Crystal structure of a superantigen bound to the high-affinity, zinc-dependent site on MHC class II. *Immunity* 14, 93-104.

Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K., Wilson, I. A. (1996) 1EBP, Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 A. *Science* 273, 464-471.

LoConte, L., Chothia, C. and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* 285, 2177-2198.

Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossman, M. G., Boege, U., Duke, G. M., Palmenberg A. C. (1987) The atomic structure of mengo virus at 3.0-A resolution. *Science* 235, 182-185.

Malby, R. L., Tulip, W. R., Harley, V. R., McKimm-Breschkin, J. L., Laver, W. G., Webster, R. G., Colman, P. M. (1994) 1NMB, The structure of a complex between the NC10 antibody and influenza virus neuraminidase and comparison with the overlapping binding site of the NC41 antibody. *Structure* 2, 733-746.

Malkowski, M. G., Martin, P. D., Guzik, J. C., Edwards, B. F. (1997) 1MKW, The co-crystal structure of unliganded bovine alpha-thrombin and prethrombin-2: movement of the Tyr-Pro-Pro-Trp segment and active site residues upon ligand binding. *Protein Sci.* 6, 1438-1448.

Malkowski, M. G., Martin, P. D., Guzik, J. C., Edwards, B. F. (1996) 1EBD, Protein-protein interactions in the pyruvate dehydrogenase multienzyme complex: dihydrolipoamide dehydrogenase complexed with the binding domain of dihydrolipoamide acetyltransferase. *Structure* 4, 277-286.

Marquart, M., Walter, J., Deisenhofer, J., Bode, W., Huber, R. (1983) 1TPA, The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Cryst B* 39, 480-487.

Mirza, O., Henriksen, A., Ipsen, H., Larsen, J., Wissenbach, M., Spangfort, M., Gajhede, M. (2000) 1FSK, Dominant epitopes and allergic cross-reactivity: complex formation between a fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen Bet V 1. *J.Immunol.* 165, 331-338.

Mittl, P. R., Di Marco, S., Fendrich, G., Pohlig, G., Heim, J., Sommerhoff, C., Fritz, H., Priestle, J. P., Grutter, M. G. (1997) 1HIA, A new structural class of serine protease inhibitors revealed by the structure of the hirustasin-kallikrein complex. *Structure* 5, 253-264.

Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Ehrlich, L. S., Summers, M. F., Carter, C., Rossmann, M. G. (1996) 1AFV, Crystal structure of dimeric HIV-1 capsid protein. *Nat. Struct. Biol.* 3, 763-770.

Monaco-Malbet, S., Berthet-Colominas, C., Novelli, A., Battai, N., Piga, N., Cheynet, V., Mallet, F., Cusack, S. (2000) 1E6J, Mutual conformational adaptations in antigen and antibody upon complex formation between an Fab and HIV-1 capsid protein P24, *Structure Fold. Des.* 8, 1069-1077.

Muller, Y. A., Chen, Y., Christinger, H. W., Li, B., Cunningham, B. C., Lowman, H. B., de Vos, A. M. (1998) 1BJ1, VEGF and the Fab fragment of a humanized neutralizing antibody: crystal structure of the complex at 2.4 A resolution and mutational analysis of the interface. *Structure* 6, 1153-1167.

Murali, R., Sharkey, D. J., Daiss, J. L., Murthy, H. M. (1998) 1BGX, Crystal structure of Taq DNA polymerase in complex with an inhibitory Fab: the Fab is directed against an intermediate in the helix-coil dynamics of the enzyme. *Proc Natl Acad Sci USA* 95, 12562-12567.

Mylvaganam, S. E., Paterson, Y., Getzoff, E. D. (1998) 1WEJ, Structural basis for the binding of an anti-cytochrome c antibody to its antigen: crystal structures of FabE8-cytochrome c complex to 1.8 A resolution and FabE8 to 2.26 A resolution. *J. Mol. Biol.* 281, 301-322.

Nassar, N., Horn, G., Herrmann, C., Block, C., Janknecht, R., Wittinghofer, A. (1996) 1GUA, Ras/Rap effector specificity determined by charge reversal. *Nat. Struct. Biol.* 3, 723-729.

Naylor, H. M., Newcomer, M. E. (1999) 1QAB, The structure of human retinol-binding protein (RBP) with its carrier protein transthyretin reveals an interaction with the carboxy terminus of RBP. *Biochemistry* 38, 2647-2653.

Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C. E., Fontecilla-Camps, J. C. (1999) 1HFE, Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure Fold. Des.* 7, 13-23.

Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K., Milburn, M. V. (1998) 1PRG, Ligand binding and co-activator assembly of the peroxisome proliferator- activated receptor-gamma, *Nature*, 395, 137-143.

Ostermeier, C., Harrenga, A., Ermler, U., Michel, H. (1997) 1AR1, Structure at 2.7 A resolution of the Paracoccus denitrificans two-subunit cytochrome c oxidase complexed with an antibody FV fragment. *Proc Natl Acad Sci USA* 94, 10547-10553, 1997.

Papageorgiou, A. C., Shapiro, R., Acharya, K. R. (1997) 1A4Y, Molecular recognition of human angiogenin by placental ribonuclease inhibitor--an X-ray crystallographic study at 2.0 A resolution. *EMBO J* 16, 5162-5177.

Parry, M. A., Fernandez-Catalan, C., Bergner, A., Huber, R., Hopfner, K., Schlott, B., Guehrs, K., Bode, W. (1998) 1BUI, The ternary microplasmin-staphylokinase-microplasmin complex is a proteinase-cofactor-substrate Complex in Action, *Nat. Struct. Biol.* 5, 917-923.

Pellegrini, L., Burke, D. F., Von Delft, F., Mulloy, B., Blu, T. L. (2000) 1E0O, Crystal Structure of Fibroblast Growth Factor Recep Ectodomain Bound to Ligand and Heparin. *Nature* 407, 1029-1034.

Perona, J. J., Tsu, C. A., McGrath, M. E., Craik, C. S., Fletterick, R. J. (1993) 1BRC, Relocating a negative charge in the binding pocket of trypsin. *J. Mol. Biol.* 230, 934-949.

Perona, J. J., Tsu, C. A., Craik, C. S., Fletterick, R. J. (1997) 1AZZ, Crystal structure of an ecotin-collagenase complex suggests a model for recognition and cleavage of the collagen triple helix. *Biochemistry* 36, 5381-5392.

Pettit, F. K. and Bowie, J. U. (1999) Protein surface roughness and small molecular binding sites. *J. Mol. Biol.* 285, 1377-1382.

Plotnikov, A. N., Schlessinger, J., Hubbard, S.R., Mohammadi, M. (1999) 1CVS, Structural basis of Fgf receptor dimerization and activation. *Cell* 98, 641-650.

Plotnikov, A. N., Hubbard, S. R., Schlessinger, J., Mohammadi, M. (2000) 1EV2, Crystal structures of two Fgf-Fgfr complexes reveal the determinants of ligand-receptor specificity, *Cell* 101, 413-424.

Prasad, L. Waygood, E. B. Lee, J. S. Delbaere, L. T. J. (1998) 2JEL, The 2.5 A structure of the Jel42 Fab fragment/Hpr complex. *J. Mol. Biol.* 280, 829-834.

Reiser, J. B., Darnault, C., Guimezanes, A., Gregoire, C., Mosser, T., Schmitt-Verhulst, A-M., Fontecilla-Camps, J. C., Malissen, B., Housset, D., Mazza, G. (2000) 1FO0, Crystal structure of a T cell receptor bound to an allogeneic MHC molecule, *Nat. Immunol.* 1, 291-298.

Rester, U., Bode, W., Moser, M., Parry, M. A., Huber, R., Auerswald, E. (1999) 1C9T, Structure of the complex of the antistasin-type inhibitor bdellastasin with trypsin and modelling of the bdellastasin-microplasmin system. *J. Mol. Biol.* 293, 93-106.

Reuter, W., Wiegand, G., Huber, R., Than, M. E. (1999) 1B33, Structural analysis at 2.2 A of orthorhombic crystals presents the asymmetry of the allophycocyanin-linker complex, AP.LC7.8, from phycobilisomes of Mastigocladus laminosus. *Proc Natl Acad Sci USA* 96, 1363-1368.

Rittinger, K., Walker, P. A., Eccleston, J. F., Smerdon, S. J., Gamblin, S. J. (1997) 1TX4, Structure at 1.65 A of RhoA and its GTPase-activating protein in complex with a transition-state analogue. *Nature* 389, 758-762.

Robinson, R. C., Radziejewski, C., Stuart, D. I., Jones, E. Y. (1995) 1B8M, Structure of the brain-derived neurotrophic factor/neurotrophin 3 heterodimer. *Biochemistry* 34, 4139-4146.

Rose, R. B., Bayle, J. H., Endrizzi, J. A., Cronk, J. D., Crabtree, G. R., Alber, T. (2000) 1F93, Structural basis of dimerization, coactivator recognition and mody3 mutations in HNF-1A, *Nat. Struct. Biol.* 7, 744-748.

Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D., Pavletich, N. P. (1998) 1BI8, Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature* 395, 237-243.

Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D., Pavletich, N. P. (1995) 1FCC, Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. Structure 3, 265-278.

Savva, R. and Pearl, L. H. (1995) 1UDI, Nucleotide mimicry in the crystal structure of the uracil-DNA glycosylase-uracil glycosylase inhibitor protein complex. *Nat. Struct. Biol.* 2, 752-757.

Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A., Kossiakoff, A. A. (1997) 1CBW, Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzheimer\'s amyloid beta-protein precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): engineering of inhibitors with altered specificities. Protein Sci. 6, 1806-1824.

Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A., Kossiakoff, A. A. (1997) 1CA0, Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzheimer\'s amyloid beta-protein precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): engineering of inhibitors with altered specificities. *Protein Sci.* 6, 1806-1824.

Schindelin, H., Kisker, C., Schlessman, J. L., Howard, J. B., Rees, D. C. (1997) 1N2C, Structure of ADP x AIF4(-)-stabilized nitrogenase complex and its implications for signal transduction. *Nature* 387, 370-376.

Schlessinger, J., Plotnikov, A. N., Ibrahimi, O.A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., Linhardt, R. J., Mohammadi, M. (2000) 1FQ9, Crystal structure of a ternary Fgf-Fgfr-heparin complex reveals a dual role for heparin in Fgfr binding and dimerization, *Mol. Cell,* 6, 743-750.

Smith, T. J., Chase, E. S., Schmidt, T. J., Olson, N. H., Baker, T. S. (1996) 1RVF, Neutralizing antibody to human rhinovirus 14 penetrates the receptor-binding canyon. *Nature* 383, 350-354.

Sogabe, S., Stuart, F., Henke, C., Bridges, A.,Williams, G., Birch, A., Winkler, F. K., Robinson, J. A. (1997) 1JRH, Neutralizing epitopes on the extracellular interferon gamma receptor (IFNgammaR) alpha-chain characterized by homolog scanning mutagenesis and X-ray crystal structure of the A6 fab-IFNgammaR1-108 complex. *J. Mol. Biol.* 273, 882-897.

Sondermann, P., Huber, R., Oosthuizen, V., Jacob, U. (2000) 1E4K, The 3.2A crystal structure of the human Igg1 Fc fragment-Fc-Gamma-Riii complex. *Nature* 406, 267-273.

Song, H. K., Suh, S. W. (1998) 1AVW, Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from Erythrina caffra and tissue-type plasminogen activator. *J. Mol. Biol.* 275, 347-363.

Starovasnik, M. A., Christinger, H. W., Wiesmann, C., Champe, M. A., De Vos, A. M., Skelton, N. J. (1999) 1QTY, Solution structure of the Vegf-Binding domain of Flt-1: comparison of its free and bound states, *J. Mol. Biol.* 293, 531-544.

Stebbins, C. E., , W. G. Kaelin Jr., Pavletich, N. P. (1999) 1VCB, Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. *Science* 284, 455-461.

Stewart, M., Kent, H. M., McCoy, A. J. (1998) 1A2K, Structural basis for molecular recognition between nuclear transport factor 2 (NTF2) and the GDP-bound form of the Ras-family GTPase Ran. *J. Mol. Biol.* 277, 635-646.

Stites, W. E. (1997) Protein-protein interactions: interface structure, binding thermodynamics, and mutational analysis. Chem. Rev. 97, 1233-1250.

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., Turk, V.(1990) 1STF, The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J* 9, 1939-1947.

Sundstrom, M., Lundqvist, T., Rodin, J., Giebel, L. B., Milligan, D., Norstedt, G. (1996) 1HWG, Crystal structure of an antagonist mutant of human growth hormone, G120R, in complex with its receptor at 2.9 A resolution. *J. Biol. Chem.* 271, 32197-32203.

Syed, R. S., Reid, S. W., Li, C., Cheetham, J. C., Aoki, K. H., Liu, B., Zhan, H., Osslund, T. D., Chirino, A. J., Zhang, J., Finer-Moore, J., Elliott, S., Sitney, K., Katz, B. A., Matthews, D. J., Wendoloski, J. J., Egrie, J., Stroud, R. M. (1998) 1EER, Efficiency of signalling through cytokine receptors depends critically on receptor orientation. *Nature* 395, 511-516.

Tan, S., Richmond, T. J. (1998) 1MNM, Crystal structure of the yeast MATalpha2/MCM1/DNA ternary complex. *Nature* 391, 660-666.

Tegoni, M., Spinelli, S., Verhoeyen, M., Davis, P., Cambillau, C. (1999) 1QFW, Crystal structure of a ternary complex between human chorionic gonadotropin (hcg) and two Fv fragments specific for the alpha and beta subunits. *J. Mol. Biol.* 289, 1375-1385.

Tegoni, M., Spinelli, S., Verhoeyen, M., Davis, P., Cambillau, C. (1997) 1AGR, Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis. *Cell* 89, 251-261.

Tesmer, J. J., Sunahara, R. K., Gilman, A. G., Sprang, S. R. (1997) 1AZS, Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS. *Science* 278, 1907-1916.

Thiel, D. J., Le Du, M.-H., Walter, R. L., D'Arcy, A., Chene, C., Fountoulakis, M., Garotta, G., Winkler, F. K., Ealick, S. E., (2000) 1FG9, Observation of an unexpected

third receptor molecule in the crystal structure of human interferon-gamma receptor complex. *Structure Fold Des.* 8, 927-936.

Tormo, J., Natarajan, K., Margulies, D. H., Mariuzza, R. A. (1999) 1QO3, Crystal structure of a lectin-like natural killer cell receptor bound to its MHC class I ligand, *Nature* 402, 623-631.

Tsai,C. J., Lin, S. L., Wolfson, H. J., Nussinov, R. (1996) A dataset of protein-protein interfaces generated with a sequence-order-independent comparision technique.*J. Mol. Biol.* 260, 604-620.

Tsunemi, M., Matsuura, Y., Sakakibara, S., Katsube, Y. (1996) 1FLE, Crystal structure of an elastase-specific inhibitor elafin complexed with porcine pancreatic elastase determined at 1.9 A resolution. *Biochemistry* 35, 11570-11576.

Tulip, W. R., Varghese, J. N., Laver, W. G., Webster, R. G., Colman, P. M. (1992) 1NCA, Refined crystal structure of the influenza virus N9 neuraminidase-NC41 Fab complex. *J. Mol. Biol.* 227, 122-148.

Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C., Watanabe, K., Nakamura, S., Takahashi, K. (1986) 1TAB, Structure of the trypsin-binding domain of Bowman Birk type protease inhibitor and its interaction with trypsin. *J. Biochem.* (Tokyo), 100, 1637-1646.

van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kroger, B., Hoffken, W., Bode, W. (1995) 1TBQ, Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. *EMBO J* 14, 5149-5157.

van de Locht, A., Stubbs, M. T., Bode, W., Friedrich, T., Bollschweiler, C., Hoffken, W., Huber, R. (1996) 1TOC, The ornithodorin-thrombin crystal structure, a key to the TAP enigma? *EMBO J* 15, 6011-6017.

van de Locht, A., Bode, W., Huber, R., Le Bonniec, B. F., Stone, S. R., Esmon, C. T., Stubbs, M. T. (1997) 1BTH, The thrombin E192Q-BPTI complex reveals gross structural rearrangements: implications for the interaction with antithrombin and thrombomodulin. *EMBO J* 16, 2977-2984.

Wang, J., Lim, K., Smolyar, A., Teng, M., Liu, J., Tse, A. G., Liu, J., Hussey, R. E., Chishti, Y., Thomson, C. T., Sweet, R. M., Nathenson, S. G., Chang, H. C., Sacchettini, J. C., Reinherz, E. L. (1998) 1NFD, Atomic structure of an alphabeta T cell receptor (TCR) heterodimer in complex with an anti-TCR fab fragment derived from a mitogenic antibody. *EMBO J* 17, 10-26.

Wang, J.-H., Smolyar, A., Tan, K., Liu, J.-H., Kim, M., Sun, Z. J., Wagner, G., Reinherz, E. L. (1999) 1QA9, Structure of a heterophilic adhesion complex between the human Cd2 and Cd58 (Lfa-3) counterreceptors. *Cell* 97, 791-803.

Wang, X., Lin, X., Loy, J. A., Tang, J., Zhang, X. C. (1998) 1BML, Crystal structure of the catalytic domain of human plasmin complexed with streptokinase, *Science* 281, 1662-1665.

Wang, Y., Jiang, Y., Meyering-Voss, M., Sprinzl, M., Sigler, P. B. (1997) 1AIP, Crystal structure of the EF-Tu.EF-Ts complex from Thermus thermophilus. *Nat. Struct. Biol.* 4, 650-656.

Waugh, S. M., Harris, J. L., Fletterick, R. J., Craik, C. S. (2000) 1FI8, The structure of the pro-apoptotic protease granzyme b reveals the molecular determinants of its specificity. *Nat. Struct. Biol.* 7, 758-764.

Wei, A., Alexander, R. S., Duke, J., Ross, H., Rosenfeld, S. A., Chang, C. H. (1998) 1KIG, Unexpected binding mode of tick anticoagulant peptide complexed to bovine factor Xa. *J. Mol. Biol.* 283, 147-154.

Welch, M., Chinardet, N., Mourey, L., Birck, C., Samama, J. P. (1998) 1A0O, Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. *Nat. Struct. Biol.* 5, 25-29.

Wiesmann, C., Ultsch, M. H., Bass, S. H., De Vos, A. M. (1999) 1WWW, Crystal Structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor receptor. *Nature* 401, 184-188.

Wu, G., Chen, Y. G., Ozdamar, B., Gyuricza, C. A., Chong, P. A., Wrana, J. L., Massague, J., Shi, Y. (2000) 1DEV, Structural basis of Smad2 recruitment by the Smad anchor for receptor activation. *Science* 287, 92-97.

Wu, G., Chai, J., Suber, T. L., Wu, J.-W., Du, C., Wang, X., Shi, Y. (2000) 1G73, Structural Basis of Iap Recognition by Smac/Diablo. *Nature* 408, 1008-1012.

Vallee, F., Kadziola, A., Bourne, Y., Juy, M., Rodenburg, K. W., Svensson, B., Haser, R. (1998) 1AVA, Barley alpha-amylase bound to its endogenous protein inhibitor BASI: crystal structure of the complex at 1.9 A resolution. *Structure* 6, 649-659.

Vallee, F., Kadziola, A., Bourne, Y., Juy, M., Rodenburg, K. W., Svensson, B., Haser, R. (1999) 1RRP, Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. *Nature* 398, 39-46.

Xiao, T., Towb, P., Wasserman, S. A., Sprang, S. R. (1999) 1D2Z, Three-dimensional structure of a complex between the death domains of Pelle and Tube. *Cell* 99, 545-555.

Zheng, N., Wang, P., Jeffrey, P. D., Pavletich, N. P. (2000) 1FBV, Structure of a C-Cbl-Ubch7 complex: ring domain function in ubiquitin-protein ligases. *Cell* 102, 533-539.

Zuccola, H. J., Filman, D. J., Coen, D. M., Hogle, J. M. (2000) 1DML, The crystal structure of an unusual processivity factor, Hsv Ul42, bound to the c terminus of its cognate polymerase. *Mol. Cell* 5, 267-278.

Ni, C.-Z., Welsh, K., Leo, E., Chiou, C.-K., Wu, H., Reed, J. C., Ely, K. R. (2000) 1FLL, Molecular basis for Cd40 signaling mediated by Traf3, *Proc. Nat. Acad. Sci.USA* 97, 10395-10399.

Chapter 6:

Summary

This chapter discusses the conclusions of the work described above, the related work of my collaborators in the lab, the relationship between the various efforts, and the planned future experiments. The studies of FcRn function were conducted in collaboration with Dr. T. S. Ramalingham, and lately, Scott Detmer, a rotation student in the lab. My contribution to these studies was to generate and purify the DNA necessary to construct the various cell lines and design, generate and purify many of the various proteins involved in the assays. This chapter discusses the implications of the FcRn:hdFc structure, implications that lead us to design an antibody with a longer serum half life. Pavel Strop of the Rees lab was instrumental in the development of this project. After describing how understanding protein-protein interactions played a role in the progress of this work, chapter concludes by discussing the new experiments to be pursued in extending these results.

Introduction

This thesis involves two distinct efforts. The first effort was an attempt to understand the functional significance of an oligomeric ribbon structure observed in FcRn/Fc complex crystals (See chapters 2, 3 and Figure 1 here for a description of the oligomeric ribbon). The basic strategy was to disrupt the various protein-protein interactions that are required to form the ribbon and observe their effect on the function of FcRn. This strategy resulted in the creation and characterization of the heterodimeric Fc (hdFc). This basic strategy also guided attempts to generate a form of FcRn that does not form an FcRn dimer. Studies of FcRn dimer formation have lead to experiments proposed at the end of this chapter that may address the significance of the oligomeric ribbon with a new technique.

The decision to study the role of the oligomeric ribbon in FcRn function by disrupting the protein-protein interactions in the ribbon lead to the second effort of this thesis. In order to disrupt the various protein-protein interactions it became important to understand the phenomenon of protein-protein interactions thoroughly. A deeper understanding of protein-protein interactions resulted in proposing experiments to improve the FcRn:Fc interaction discussed at the end of this chapter. The deeper understanding of protein-protein interactions has also indicated possible difficulties in the strategy of disrupting the oligomeric ribbon. Circumventing these difficulties is possible in the experiments utilizing the new strategy mentioned above.

The oligomeric ribbon
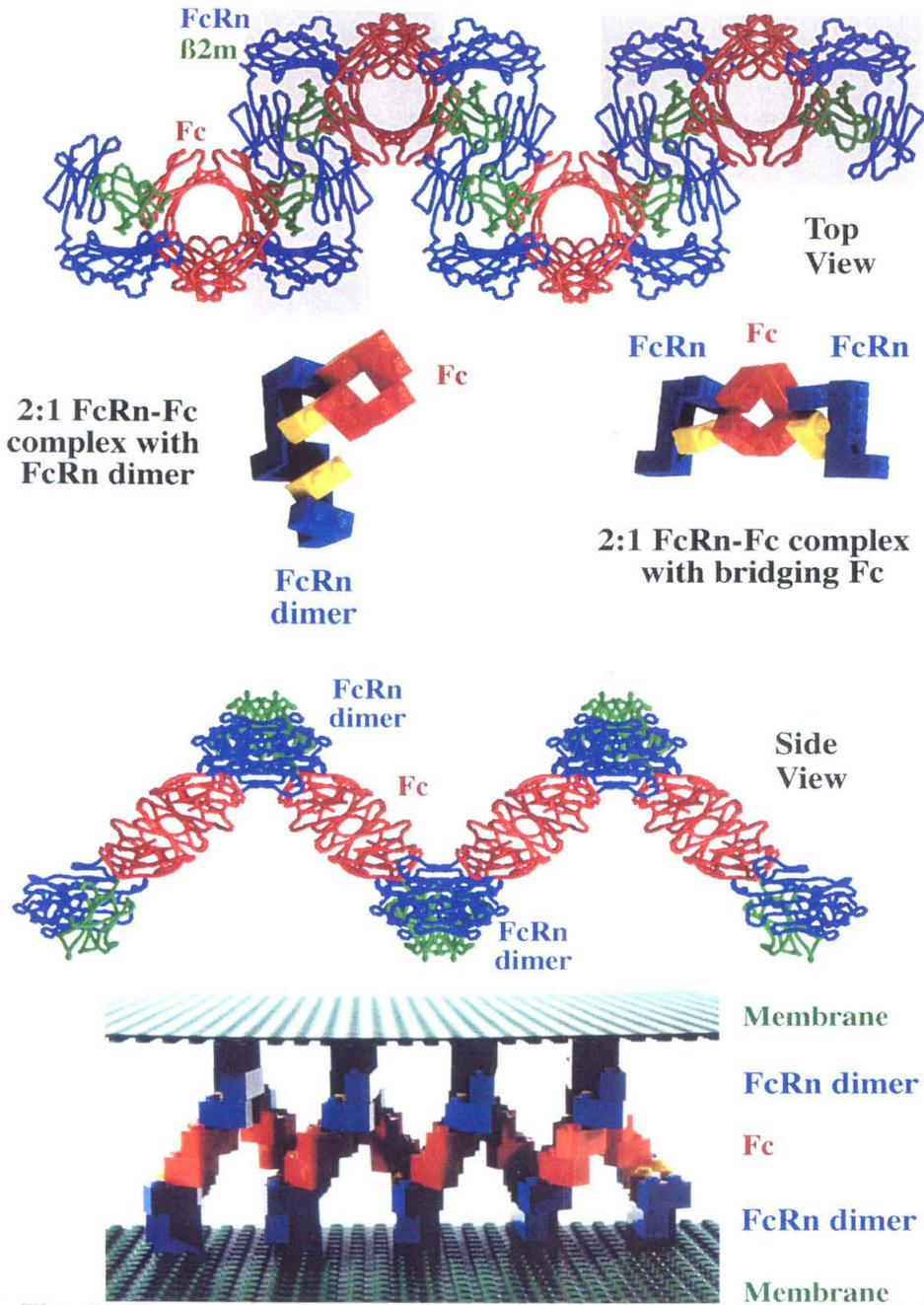
Figure 1: The oligomeric ribbon observed in the FcRn-Fc cocrystals as a model for the formation of higher order oligomers upon binding of IgG to membrane bound FcRn. A) Top view of the ribbon. FcRn dimers are bridged by Fc molecules. Under physiological conditions, each FcRn dimer would be associated with a membrane parallel to the plane of the paper: the left-most dimer is associated with a membrane below the paper; the central dimer with a membrane above the paper, and the right-most dimer is again associated with a membrane below the paper. Lego representations of the shaded portions are shown below the corresponding region of the ribbon. B) Side view, rotated by 90° about the horizontal axis from the view in part A. The FcRn dimers are seen looking down their long axes (vertical in A). Formation of the ribbon in acidic vesicles would result in close juxtaposition (~200 Å) of two parallel membranes and a clustering of FcRn cytoplasmic tails into an ordered array.

Discussion:

Attempts to determine the importance of the "oligomeric ribbon" to FcRn function

The possibility that the oligomeric ribbon that forms in FcRn/Fc complex crystals is important to FcRn function stimulated the first effort of my thesis. We proposed that the oligomeric ribbon might be important to FcRn function if it formed inside cells between narrowly opposed (~200 Å) membranes (Chapter 2). In MDCK cells expressing FcRn or FcRn-EGFP, FcRn is localized in apical endosomes, where the acidic pH permits FcRn-IgG binding (Ramalingham et al., manuscript in preparation). From the apical endosomes of these cells FcRn/IgG complexes are either trancytosed to basolateral surface of the cells or recycled to the cell surface (Ramalingham et al., manuscript in preparation). We had several conjectures as to when the oligomeric ribbon might occur during these processes. Small FcRn-IgG complexes with all the interactions seen in the oligomeric ribbon might form as a signal for transport inside endosomes. More extensive ribbon-like structures might also exist as narrow tubules transiently extending from endosomes to the cell surface. The narrow size of the oligomeric ribbon and the its possibly ephemeral *Nature* made it unclear whether visualizing the ribbon directly was possible. Even if the oligomeric ribbon could be visualized directly it would still not be clear whether forming the ribbon was necessary for FcRn to recycle or transcytose IgG.

Ribbon testing strategy

In order to determine whether ribbion formation was necessary to FcRn function, we proposed to disrupt the ribbon and determine if FcRn would still trancytose IgG. The most direct way to disrupt the ribbon was to alter its components. The formation of the oligomeric ribbon requires two distinct 2:1 FcRn:Fc complexes: a 2:1 complex with an

FcRn dimer and a2:1 complex with a bridging Fc. We thought that if we could disrupt each of these complexes selectively and if each of these disruptions to the ribbon were deleterious to FcRn function we could argue persuasively that the ribbon was necessary for FcRn function.

The earliest experiments in which we participated, indicated that one of the 2:1 FcRn:Fc complexes, the complex of one Fc bound to a dimer of FcRn molecules, was important to FcRn function. In the experiments described in Appendix A, we found that mutations made at positions in the FcRn dimer interface affected the affinity of the interaction between FcRn and IgG. The apparent importance of the FcRn dimer to the high affinity binding of IgG lead us to determine whether disrupting the other 2:1 FcRn:Fc complex would affect FcRn function.

Generating and characterizing the hdFc

We were successful in disrupting the other 2:1 FcRn:Fc complex, the complex of wild-type Fc (wtFc) bound on either side by FcRn. This complex was disrupted by the generation of the hdFc (Chapter 2). The hdFc is a wild-type Fc molecule with one FcRn binding site removed by site-directed mutagenesis. In solution studies and in experiments with FcRn anchored to a solid support, the hdFc does not bind a second FcRn (Chapter 2). We also showed that hdFc binds FcRn as well as wtFc binds FcRn ats high affinity binding site but that the wtFc shows higher avidity for the ligand. Because the hdFc bound one FcRn with high affinity we felt that examining its behavior in a trancytosis assay would indicate whether the complex of Fc bound on either side by FcRn was necessary for FcRn function.

hdFc in functional assays

The disruption of the complex of Fc bound on either side by FcRn by the hdFc was shown to be deleterious to FcRn function (Ramalingham et al., manuscript in preparation, Figure 2). The hdFc was trancytosed at levels not much above background. Ramalingham also developed a method for assaying the formation of ribbon like structure in a test tube. In an assay using liposomes containing a GPI-linked form of FcRn, the wtFc induced the aggregation of liposomes at pH 6 but not at pH 8 (Ramalingham et al., manuscript in preparation, Figure 3). The hdFc largely fails to induce liposome aggregation in this assay. It is not clear what causes the small level of trancystosis of hdFc or the slight tendency of the hdFc to induce liposome aggregation. As the wtFc is expressed in the same cells as the hdFc these small effects might be the result of residual amounts of wtFc in the purified hdFc. Because this small level of activity occurs in both assays we believe that the small level of hdFc that undergoes trancytosis is a part of experimental noise and that is not inconsistent with the oligomeric ribbon being necessary for FcRn function.

hdFc in FcRn:hdFc cocrystals

The hdFc forms FcRn:hdFc crystals that do not contain the oligomeric ribbon. These crystals diffracted to 2.8 Åresolutions and allowed the first detailed analysis of the FcRn/Fc interface (Chapter 3). The structure contains ordered carbohydrate at the FcRn/Fc interface that was helpful in explaining why high mannose forms of FcRn bound Fc less well than FcRn forms with complex carbohydrates (Popov et al., 1996; Sanchez et al., 1999). The FcRn:hdFc structure illustrated the mechanism of pH-dependent binding. Three histidines form salt-bridges with 2 glutamates and one aspartate. The one salt-
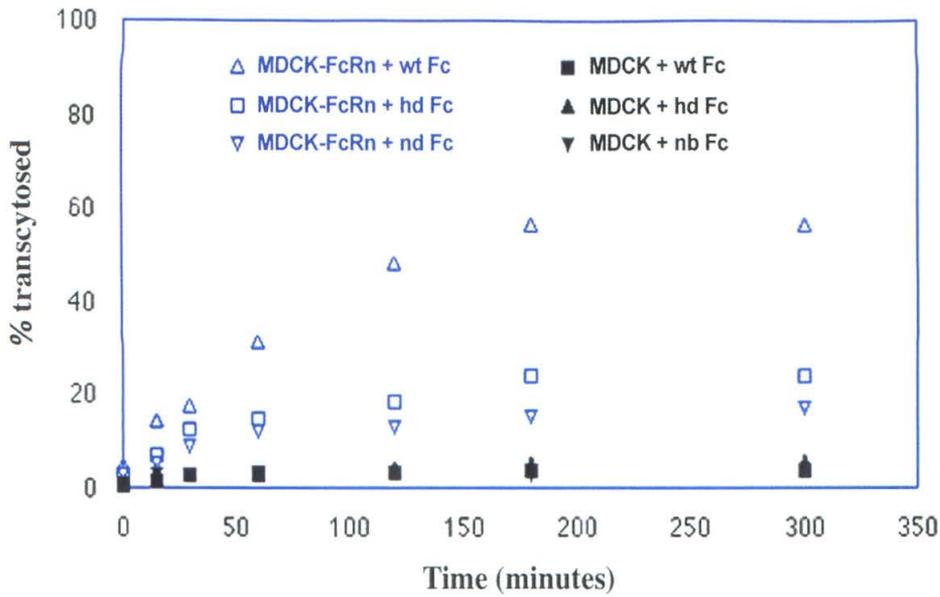
Figure 2: Transcytosis of various Fcs in MDCK cells stably transfected with FcRn.
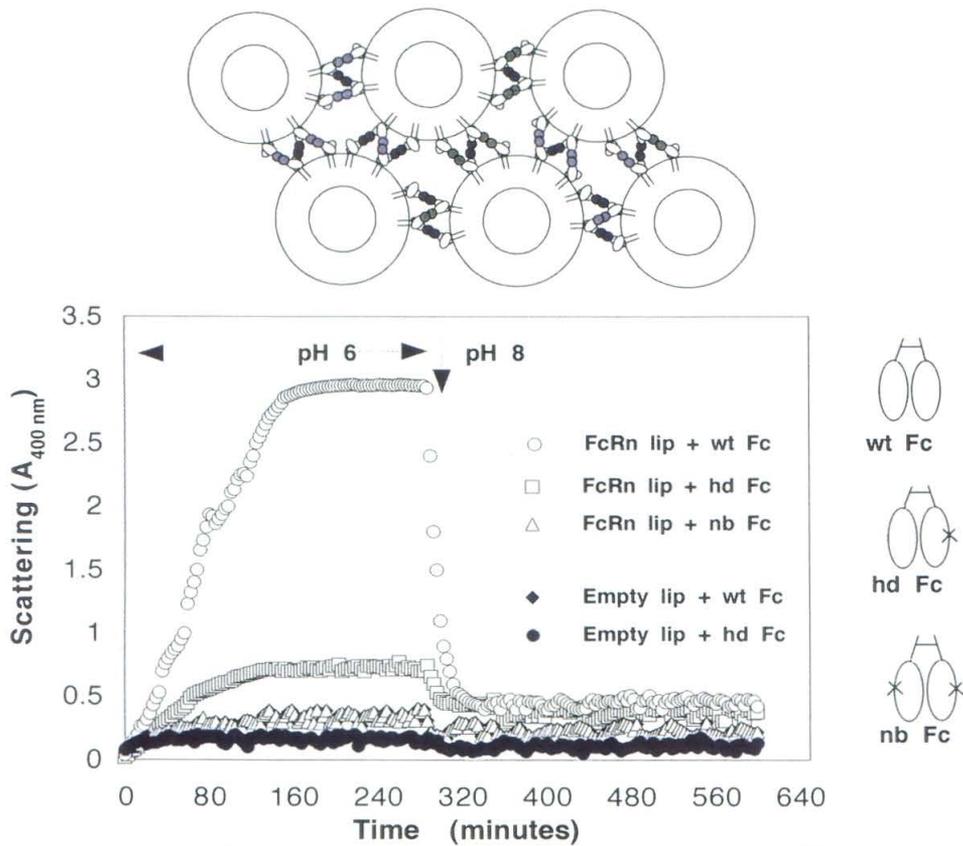
Figure 3: Aggregation of liposomes containing GPI-linked FcRn by wtFc and not by hdFc.

bridge that is conserved in the rat, mouse, bovine, and human sequences of FcRn and IgG is the salt-bridge that is most buried in the structure. In conjunction with the extensive mutagenesis work done by our lab and others we were able to use this structure to suggest strategies for making IgG antibodies with longer serum half-lives. One of these strategies, an attempt to bias the observed flexibility between the $C\gamma2$ and $C\gamma3$ domains (proposed in Chapter 3) towards the relative orientation between the domains observed when binding FcRn is a future experiment we intend to pursue.

FcRn dimer studies

The hdFc gave us the ability to disrupt the complex of Fc bound on either side by FcRn and the ability to distinguish between this complex and 2:1 FcRn:Fc complex involving the formation of the FcRn dimer (Chapter 4). We used this ability to show that the FcRn-TrkA chimera in the presence of hdFc was capable of forming the FcRn dimer. It was our hope that by introducing mutations at positions in the FcRn dimer interface to the FcRn-TrkA chimera we would be able to isolate a form of FcRn that does not undergo ligand-induced dimerization. Unfortunately, by the time these mutant forms were prepared we were no longer able to observe the hdFc inducing the FcRn dimer. We are unsure why the results of this experiment changed but one possibility is an alteration of the cell line, perhaps in response to non-native signalling by the chimeric protein. Because of the failure to reproduce this result and because the hdFc did not induce a dimer of FcRn molecules in the FcRn:hdFc crystals (Chapter 3) we designed an assay that would allow us to determine if FcRn forms a dimer without inducing any signalling cascades and independent of the presence of the ligand.

FcRn dimer FRET assays

We designed FcRn-ECFP and FcRn-EYFP chimeras, transiently transfected Cos-7 cells with the constructs and assayed for fluorescence resonance energy transfer (FRET) using a two-photon microscope (Chapter 4). We were able to observe FRET in some cells. These results were consistent with the formation of the FcRn dimer in these cells but they are not conclusive proof. There was difficulty in assessing whether the observed FRET was the result of the formation of the FcRn dimer or the result of general crowding in cells expressing the chimeras at high levels. This difficulty combined with the inconsistency and variability in these results suggest that more interpretable results might be obtained with cell lines stably transfected with these chimeras.

The current view of the importance of the oligomeric ribbon to FcRn function

The results of this thesis are not conclusive on the importance of the oligomeric ribbon to FcRn function. The importance of residues at the FcRn dimer interface for Fc binding is consistent with an FcRn dimer being used in cells to bind Fc. The low amount of hdFc trancytosed by FcRn is consistent with both FcRn binding sites being required for efficient transcytosis. The absence of the FcRn dimer in the FcRn:hdFc crystals and the difficulty in observing the formation of the FcRn dimer in conditions outside the crystals does not demonstrate the FcRn dimer is unimportant but does not support its importance either. The wtFc-dependent aggregation of liposomes is consistent with oligomeric ribbon structures forming between the liposomes but it is unclear whether these studies replicate what occurs inside the cell.

Lessons gleaned from reviewing protein-protein interactions

Our study of protein-protein interactions began early in this work and was the result of wishing to know how best to disrupt the FcRn:Fc interaction. We learned that not all residues are equally represented in nonobligate protein-protein interfaces and that not all residues in nonobligate protein-protein interfaces are equally important to the formation of the complex (Chapters 1 and 5). In particular, we learned that only the residues at the core of these interfaces are likely to be critical to the formation of the complex. We also learned that these residues are usually buried away from solvent and extremely well packed. This latter understanding was helpful in deciding how to try to enhance the FcRn:Fc interaction (see below). Finally, reviewing protein-protein interactions revealed that while disrupting a high affinity interaction is easy to detect, disrupting interactions of low affinity is much more difficult. Because the affinity of the FcRn-FcRn dimer interaction is on the order of millimolar (Chapter 4) even in the presence of the ligand (Chapter 3), disrupting the formation of this complex will be difficult to detect. This illuminates how one aspect of our general strategy for demonstrating the ribbon was impractical. Because disrupting the FcRn dimer is difficult to detect, we feel that demonstrating its importance to FcRn function might be better accomplished by determining when and where it occurs as FcRn transcytoses IgG.

Proposed experiments

It is the conclusion of this thesis that two experiments in this system should be pursued. To monitor the formation of the FcRn dimer throughout the FcRn-mediated trancystosis of IgG we propose to generate stable FcRn-ECFP/FcRn-EYFP transfectant of MDCK cells as proposed in chapter 4. This cell line will allow us to monitor the

formation of FcRn dimers directly by two-photon microscopy and it will allow us to circumvent obtaining mutant forms of FcRn that do not form an FcRn dimer.

The second experiment we propose is to generate a mutant form of the Fc protein that may have a higher affinity for FcRn. For reasons outlined in Chapter 3 we propose to make a high affinity Fc by constraining the considerable flexibility between the Cγ2 and Cγ3 domains. To constrain this flexibility we decided to try pack this interface as densely as possible. To do this we collaborated with Pavel Strop of the Rees group who utilized the ORBIT program (Dahiyat and Mayo, 1997) to predict amino acid substitutions that would generate a stable core between these two domains. To do this the amino acids at positions in between the Cγ2 and Cγ3 domains were selected by visual inspection. The type of amino acids and the conformation of the side chain at these positions were allowed to vary. At all other positions the amino acid type was held constant but the conformation of the side chain could vary. The positions of the main chain atoms were held constant for the calculations. The results of the calculations are summarized below.

Positions	247	248	250	251	314	338	376	428	430
Wild-type	Thr	Lys	Val	Leu	Leu	Lys	Asp	Leu	Glu
Mutant	Val	Arg	Ile	Leu	Leu	Lys	Leu	Phe	Glu

Independent calculations by the author confirm that the density of atoms in the Cγ2 and Cγ3 interface should be significantly improved in the mutant protein. Generating this mutant protein and assaying its ability to bind FcRn will allow us to

determine if packing the Cγ2 and Cγ3 interface results in an Fc that binds FcRn better. Obtaining crystals of the mutant protein will allow us to determine if packing the Cγ2 and Cγ3 interface biased the orientation of the two domains towards the conformation they assume when binding FcRn.

References

Dahiyat, B. I. and Mayo. S. L. (1997) De novo protein design: fully automated sequence selection. *Science*. 278, 82-87.

Popov, S., Hubbard, J. G., Kim, J. –K., Ober, B., Ghetie, V., and Ward, E. S. (1996) The stoichiometry and affinity of interaction of murine Fc fragments with the MHC class I-related receptor, FcRn. *Mol. Immunol.* 33, 521-530.

Ramalingham, T. S., Detmer, S., Martin, W. L., and Bjorkman, P. J. Mechanism of IgG recycling and transcytosis by the neonatal Fc receptor expressed in MDCK cells. Manuscript in preparation.

Sanchez, L. M., Penny, D. M., and Bjorkman, P. J. (1999) Stoichiometry of the interaction between the MHC-related Fc receptor and its Fc ligand. *Biochemistry* 38, 9471-9476.

Appendix A:

Identification of Critical IgG Binding Epitopes on the

Neonatal Fc Receptor

This appendix describes the characterization of the Fc binding site on FcRn and the FcRn dimer interface by assaying the ability of mutant forms of FcRn to bind IgG in surface plasmon resonance experiments. My contribution to this work occurred during my rotation in the Bjorkman lab and involved making six of the 11 site-directed mutations by the Kunkel method. I also subcloned the DNA for the mutant proteins into the mammalian cell expression vectors.

JMB

Identification of Critical IgG Binding Epitopes on the Neonatal Fc Receptor

Daniel E. Vaughn[1], Christina M. Milburn[1,2], David M. Penny[1,2] W. Lance Martin[1], Jennifer L. Johnson[1,2] and Pamela J. Bjorkman[1,2]*

[1]*Division of Biology 156-29 and* [2]*Howard Hughes Medical Institute, California Institute of Technology, Pasadena CA 91125, USA*

The neonatal Fc receptor (FcRn) binds maternal immunoglobulin G (IgG) during the acquisition of passive immunity by the fetus or newborn. FcRn also binds IgG and returns it to the bloodstream, thus protecting IgG from a default degradative pathway. Biosensor assays have been used to characterize the interaction of a soluble form of rat FcRn with IgG, and demonstrate that FcRn dimerization and immobilization are necessary to reproduce *in vivo* binding characteristics. Here, we report the identification of several FcRn amino acid substitutions that disrupt its affinity for IgG and examine the effect of alteration of residues at the FcRn dimer interface. The role of these amino acids is discussed in the context of the previously reported structures of rat FcRn and a complex of FcRn with the Fc portion of IgG.

© 1997 Academic Press Limited

Keywords: surface plasmon resonance; biosensor; structural binding site; functional epitope; binding free energy

*Corresponding author

Introduction

The neonatal Fc receptor (FcRn) binds immunoglobulin G (IgG) in two important physiological processes (Junghans, 1997). In the transmission of passive immunity from mother to offspring, FcRn mediates the transcytosis of maternal IgG across fetal and/or neonatal tissues, depending on the species involved. FcRn also binds IgG inside cells that degrade serum proteins and returns it to the bloodstream, thus rescuing it from a default degradative pathway. Each of these processes is facilitated by the pH dependence of FcRn interaction with IgG (Rodewald, 1976; Rodewald & Kraehenbuhl, 1984). Under slightly acidic con-

ditions (pH 6.0), FcRn binds IgG with a K_D of approximately 20 nM, while under slightly alkaline conditions (pH 7.5) there is no detectable binding for IgG concentrations of several μM (Raghavan *et al.*, 1995a).

Unlike other known Fc receptors which are composed of tandemly repeated immunoglobulin-like domains, FcRn is homologous to class I major histocompatibility complex (MHC) molecules (reviewed by Raghavan & Bjorkman, 1996). The FcRn light chain is β2-microglobulin (β2m; Simister & Rees, 1985), the same light chain that associates with class I MHC heavy chains. The heavy chains of both FcRn and class I MHC molecules consist of three extracellular domains, α1, α2 and α3, followed by a transmembrane region and a short cytoplasmic sequence (Bjorkman & Parham, 1990; Simister & Mostov, 1989). The extracellular region of FcRn and class I MHC heavy chains exhibit low but significant sequence identity (22 to 30% for the α1 and α2 domains, 35 to 37% for the α3 domain; Simister & Mostov, 1989). A 2.2 Å crystal structure of the extracellular region of rat FcRn revealed the expected structural similarity to MHC molecules and suggested the participation of a dimer of FcRn heterodimers in ligand binding (Burmeister *et al.*, 1994a). In each of the three crystal forms in the FcRn structure determination, a dimer of FcRn molecules mediated by contacts between the α3 and β2m domains was observed. The FcRn dimer

© 1997 Academic Press Limited

was also observed in the crystal structure of a 2:1 complex of FcRn and the Fc fragment of rat IgG which was solved to a resolution of ≈6.5 Å (Burmeister *et al.*, 1994b). Although the limited resolution of this structure determination prohibited a detailed analysis of the molecular interaction, the approximate binding region on each molecule could be localized (Figure 1).

Results from studies to map the binding sites on FcRn and IgG are consistent with the binding interaction observed in the FcRn/Fc crystal structure. Three sites on the IgG ligand have been identified as important for the binding of mouse and rat FcRn: Ile253, His310, and the region including His433, His435 and His436 (Kim *et al.*, 1994a,b,c; Medesan *et al.*, 1997; Raghavan *et al.*, 1994, 1995a; Figure 2a and b). Mutation of residues within each of these epitopes disrupts both binding to FcRn *in vitro* and FcRn mediated transport *in vivo* (Kim *et al.*, 1994a,b,c; Medesan *et al.*, 1997; Raghavan *et al.*, 1995a). There is less information available concerning the functional epitope on FcRn for binding IgG. Site-directed mutagenesis was used to identify two regions on rat FcRn that exert a slight effect on its affinity for rat IgG: His 250 and His251, and the α3 domain loop comprising resi-

dues 219 to 224 (Raghavan *et al.*, 1994). Neither of these regions is at the direct interface between the "primary" FcRn molecule that shows the majority of contacts with Fc in the 2:1 complex involving the FcRn dimer (Figure 1). His250 and His251 are at the FcRn dimer interface where they may exert their effect in IgG binding affinity either by modulating formation of the FcRn dimer, which is required for high affinity binding of IgG (Raghavan *et al.*, 1994, 1995b), or by direct contact between Fc and the "secondary" FcRn molecule of the FcRn dimer (Figure 1; Z. Weng, K. Gulukota, D.E.V., P.J.B. & C. DeLisi, unpublished results). Similarly, residues 219 to 224 of the secondary, but not the primary, FcRn molecule are in a position to contact Fc (Z. Weng, K. Gulukota, D.E.V., P.J.B. & C. DeLisi, unpublished results).

Because of the low resolution of the available FcRn/Fc co-crystal structure, we cannot identify specific FcRn residues at the IgG interface with certainty. To characterize further the epitope on FcRn, we designed, constructed, and expressed mutants of rat FcRn at the predicted interfaces with IgG and with the second FcRn in the receptor dimer. We have quantitatively characterized the IgG binding interaction for each of these FcRn mutants using a biosensor based assay (Raghavan *et al.*, 1994, 1995a,b; Vaughn & Bjorkman, 1997). In this assay, soluble FcRn immobilized on a biosensor surface reproduces the affinity for IgG and the pH dependent interaction observed for membrane bound FcRn *in vivo* and on transfected cells (Mackenzie, 1984; Raghavan *et al.*, 1994). We have used the assay to demonstrate that high affinity binding of IgG requires both FcRn dimerization and receptor immobilization (Raghavan *et al.*, 1995b; Vaughn & Bjorkman, 1997). Here, we identify several residues on rat FcRn that exert a major effect on binding affinity for murine or rat IgG, discuss the implications for pH dependent FcRn/IgG binding and FcRn dimerization, and compare the results to other known receptor–ligand and protein–protein interaction systems.

Results

β2m mutants at the FcRn–Fc interface

Using the FcRn/Fc co-crystal structure (Burmeister *et al.*, 1994b), we identified FcRn residues predicted to contact Fc. Most are located on the α2 domain of the primary receptor, with fewer contributions from the α1 and β2m domains of the primary receptor and the α3 domain of the secondary (dimer-related) receptor. To verify the contacts between Fc and the primary FcRn predicted by the low resolution co-crystal structure, we constructed eight FcRn mutants.

The co-crystal structure predicts that the N-terminal portion of β2m contacts the Fc C_H2 domain (Figures 1 and 2a). To test this hypothesis we examined the binding of three mutants: βI1A, βQ2A, and β-1Y (Table 1). In order to
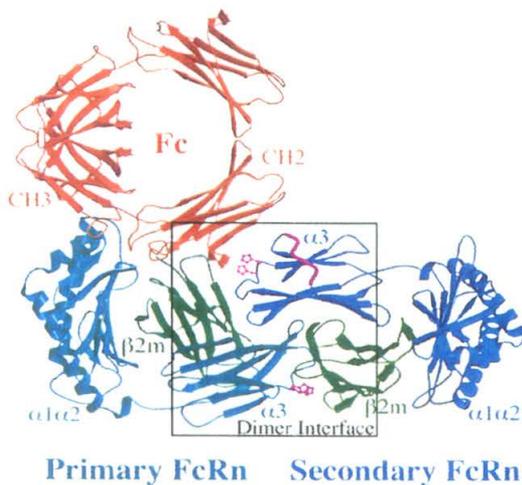
Figure 1. Ribbon diagram of the co-crystal structure of the FcRn dimer binding to its Fc ligand. The high affinity binding site for IgG is formed by a homodimer of two FcRn molecules identified as the "primary FcRn" (cyan and light green) and the "secondary FcRn" (blue and dark green). Each FcRn molecule is composed of a multidomain heavy chain (cyan or blue) and a single domain light chain (β2m; green). The Fc ligand (red) interacts with the heavy chain α2 domain and β2m domain of the primary FcRn, and the heavy chain α3 domain of the secondary receptor. Previously characterized FcRn heavy chain mutations that affect ligand binding (see text for details) are shown in magenta. The FcRn dimer interface shown as a close-up in Figure 2c is indicated by a black box.

Primary FcRn   Secondary FcRn

Figure 2. Close-up views of the mutated amino acids on FcRn. Orientations are the same as shown in Figure 1 (a and b) or rotated by 45° about a horizontal axis (c). Substituted amino acids are highlighted in yellow (mutations described and characterized in this text) or magenta (previously characterized mutations). Other colors are as described for Figure 1. Highlighted positions are: (a) Ile1 (I1) and Gln2 (Q2) of β2m, amino acids 84 to 86 and Trp133 (W133) of the FcRn heavy chain, and Fc Ile253 (I253); (b) His310 (H310), His433 (H433), His435 (H435), and His436 (H436) of the Fc and Glu117 (E117), Glu132 (E132), Glu135 (E135), and Asp137 (D137) of the FcRn heavy chain; and (c) Gly191 (G191), His250 (H250), His251 (H251) of the FcRn heavy chain, and Glu89 (E89) of β2m.

avoid potential exchange of mutant rat β2m for endogenous hamster β2m or bovine β2m in the medium, these mutations were introduced into a lipid-linked version of rat β2m that associates with the truncated FcRn heavy chain at the cell surface. This form of lipid-linked wild-type FcRn was previously shown to bind IgG with the known physiological pH dependence (Gastinel *et al.*, 1992). The affinity of these mutants for radiolabeled IgG was evaluated using a cell binding assay (Table 2). The binding response at equilibrium was plotted as a function of the log of the concentration of IgG and fit to a model assuming one class of binding sites (Figure 3a). The binding behavior of FcRn at the cell surface may be more complicated (see Vaughn & Bjorkman, 1997 for discussion), but the data from the cell binding assays are too noisy to reliably

fit more complicated binding models. However, the affinity derived using the cell binding assay agrees with high affinity binding constants obtained using a biosensor assay ($K_D = 24$ nM, cell binding assay; $K_D = 24$ nM, biosensor assay (Vaughn & Bjorkman, 1997); in both cases a rat IgG2a monoclonal antibody against CD4 was used). When the mutants were examined using the cell binding assay, we find that substitution of the N-terminal Ile of β2m with alanine (βI1A), or extension of the N terminus by one additional residue (β-1Y), eliminates significant IgG binding at concentrations up to 0.9 µM, corresponding to a $\Delta\Delta G$ of >2 kcal/mol (Table 2). However, substitution of Gln2 of the β2m domain with alanine (βQ2A) results in binding similar to wild-type FcRn (Table 2; Figure 3a and b), although the total binding response is increased presumably

Table 1. FcRn mutants

Mutant	Domain	Position(s)	Mutation
Fc–FcRn interface			
βI1A	β2m-DAF	1	I → A
βQ2A	β2m-DAF	2	Q → A
β-1Y	β2m-DAF	N terminus	add Y
W133A	α2	133	W → A
84–86	α1	84–86	NQI → GYY
E117S	α2	117	E → S
E132Q&E135Q	α2	132 and 135	Both E → Q
D137N	α2	137	D → N
Dimer interface			
G191E	α3	191	G → E
G191H	α3	191	G → H
βE89H	β2m	89	E → H
βE89K	β2m	89	E → K

Table 2. Characterization of IgG binding to FcRn mutants

Mutant	$K_{D,1}$ (nM)	f_1 (%)	$K_{D,2}$ (μM)	$R_{max,tot}$ (RU)	$\Delta\Delta G$ (kcal/mol)	Receptor (RU)
Fc–FcRn interface:						
Wild-type	24 ± 4			0.129 ± 0.008[a]		(cell)[b]
βI1A	≫900			N.B.	≫2	(cell)
βQ2A	10 ± 2			0.60 ± 0.04[a]	−0.5	(cell)
β-1Y	≫900			N.B.	≫2	(cell)
Wild-Type	12 ± 2	43 ± 2	1.1 ± 0.3	910 ± 30		2370
W133A	>8000			130[c]	>4	3370
84–86	6.0 ± 0.6	62 ± 2	0.9 ± 0.2	1430 ± 30	−0.4	1850
Wild-Type	4.9 ± 0.4	66 ± 2	0.7 ± 0.2	1890 ± 30		2660
D137N	>8000			250[c]	>4	2160
E117S	≫8000			N.B.	≫4	3820
E132Q&E135Q	≫8000			N.B.	≫4	3770
Dimer interface:						
Wild-Type	6.3 ± 0.4	60 ± 1	0.9 ± 0.1	1410 ± 20		2660
G191E	25 ± 5	41 ± 3	2.2 ± 0.7	650 ± 40	0.8	2820
G191H	21 ± 4	39 ± 2	1.9 ± 0.5	740 ± 30	0.7	3600
Wild-type	10 ± 1	54 ± 3	0.9 ± 0.2	580 ± 20		1460
G191E	57 ± 5	41 ± 1	4.4 ± 0.9	370 ± 20	1.1	2450
G191H	90 ± 30	37 ± 4	5 ± 4	350 ± 60	1.4	1650
Wild-type	6.7 ± 0.6	54 ± 2	0.8 ± 0.1	1250 ± 20		2700
βE89K	19 ± 2	43 ± 3	1.2 ± 0.2	1180 ± 30	0.6	2670
βE89H	22 ± 2	46 ± 2	1.4 ± 0.2	1310 ± 20	0.7	2000
Wild-type	7 ± 1	54 ± 3	0.6 ± 0.2	880 ± 20		1170
βE89K	27 ± 2	38 ± 2	1.7 ± 0.5	720 ± 30	0.8	1390
βE89H	22 ± 2	37 ± 3	1.5 ± 0.4	660 ± 20	0.7	1450

N.B. = no binding detected.
[a] Arbitrary units on the same relative scale.
[b] Affinity measurements done using lipid-linked FcRn at the cell surface.
[c] Observed binding response at highest IgG concentration (8 μM).

because the β2QA mutant is expressed at a higher level than the wild-type construct as demonstrated by fluorescence-activated cell sorting (FACS) analysis (data not shown).

FcRn heavy chain mutants at the FcRn–Fc interface

In order to characterize more completely the binding, we used a surface plasmon resonance assay to measure the binding affinities between IgG and soluble forms of wild-type and mutant receptors. Biosensor analyses of the interaction of

IgG with soluble FcRn immobilized on a biosensor chip show non-linear behavior in Scatchard plots (Raghavan *et al.*, 1995a,b). We recently demonstrated that the interaction of IgG with immobilized FcRn is best described as two classes of non-interacting FcRn/IgG complexes: a slow-dissociating population with an affinity in the range of values calculated in studies using membrane-bound FcRn (Mackenzie, 1984; Raghavan *et al.*, 1994), and a fast-dissociating population with lower affinity (Vaughn & Bjorkman, 1997). As previously described (Vaughn & Bjorkman, 1997), plots of the net equilibrium response *versus* the log
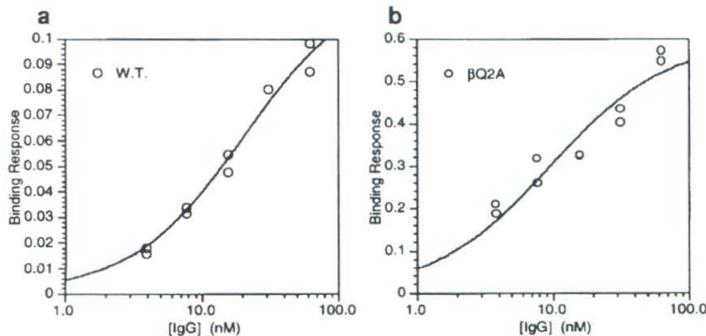
Figure 3. IgG binding to β2m mutants at the FcRn–Fc interface. Equilibrium IgG binding to membrane associated wild-type and mutant βQ2A FcRn (expressed in arbitrary units proportional to the amount of IgG bound) is plotted as a function of the log of the IgG concentration. The best fit binding curves (continuous lines), modelled as a single class of non-interacting binding sites, correspond to the values reported in Table 2. Two additional mutants βI12A and β-1Y) showed no significant binding at IgG concentrations of ≤0.9 μM (data not shown).
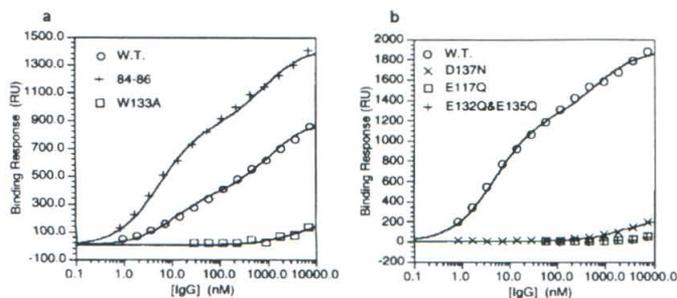
Figure 4. IgG binding to FcRn heavy chain mutants at the FcRn–Fc interface. Equilibrium IgG binding to wild type and mutant FcRn proteins immobilized on a biosensor chip is plotted as a function of the log of the IgG concentration. The best fit binding curves (continuous lines), modelled as two classes of non-interacting binding sites, correspond to the values reported in Table 2.

of the concentration of injected ligand were analyzed using non-linear regression to derive the percent of FcRn/IgG complexes comprising each population (f_1 and f_2), the high and low affinity binding constants ($K_{D,1}$ and $K_{D,2}$, respectively), and the maximum combined binding response ($R_{max,tot}$; Table 2).

Two mutations to the FcRn heavy chain were made in order to test the importance of potential aromatic and/or aliphatic contacts predicted from the co-crystal structure: W133A and 84-86 (Table 1). FcRn Trp133 is exposed to solvent in the structure of FcRn alone (Burmeister et al., 1994a), and is in a position where it could interact with the Ile253 of Fc, which was identified as a critical residue on Fc for interaction with FcRn (Kim et al., 1994a,b,c; Medesan et al., 1997; Raghavan et al., 1995a) (Figure 2a). We find that IgG binding to the W133A mutant is extremely weak even at µM concentrations (Figure 4a). Thus the affinity of this mutant for IgG is >8 µM corresponding to a $\Delta\Delta G$ of >4 kcal/mol (Table 2). The 84-86 mutation (Table 1) encompasses three of the four amino acids in the α1 domain that are predicted to be within 5 Å of the Fc ligand in the co-crystal structure (Figure 2a). We find that replacement of these residues with the corresponding residues from class I MHC molecules (Kabat et al., 1991) results in a binding affinity similar to wild-type (Figure 4a; Table 2).

Much of the pH dependence of the FcRn–Fc interaction has been assigned to the titration of histidine residues on the Fc ligand (Raghavan et al., 1995a). There are several acidic residues on the α2 domain of the receptor that could interact electrostatically with the protonated form of these histidine residues. Glu117 is positioned near Fc His310, Glu132 and Glu135 are near Fc His435, and Asp137 is in the vicinity of Fc His 435 (Figure 2b). Three mutations (E117S, E132Q&E135Q, and D137N) were constructed to test the functional importance of these acidic residues (Table 1). Neither E117S nor E132Q&E135Q bind IgG significantly at concentrations up to 8 µM (Figure 4b; Table 2). The D137N mutant binds IgG slightly at µM concentrations, but the binding affinity is also >8 µM (Figure 4b; Table 2). For each mutant, the calculated high affinity binding is reduced by

greater than 1000-fold, corresponding to a $\Delta\Delta G$ of more than 4 kcal/mol.

FcRn mutants at the dimer interface

In order to examine the effects on IgG binding of mutations at the FcRn–dimer interface, we made mutations at two additional positions within FcRn. Based on the crystal structures of FcRn alone and the FcRn/Fc complex (Burmeister et al., 1994a,b), Gly191 is positioned such that side-chains introduced at this position should interact with position 191 of the dimer related FcRn molecule (Figure 2c). We constructed two mutants (G191E and G191H) to examine the effect of substitution at position 191 (Table 1). Both mutations have a slight (≈fourfold) effect on the high affinity binding constant ($K_{D,1}$) when FcRn is coupled at densities greater than 2500 RU (Figure 5a; Table 2). Both mutants also show a reduction in the total number of binding sites ($R_{max,tot}$) and the fraction of high affinity binding sites (f_1) compared to wild-type FcRn. IgG binding is further reduced when the FcRn concentration on the chip is lowered by coupling at lower densities. Under these conditions, high affinity binding is reduced (six- to tenfold) and the fraction of high affinity and total number of binding sites are again reduced relative to wild-type FcRn (Figure 5b; Table 2).

We next examined the effects of mutating β2m residue Glu89. This residue is in position to form a salt bridge with either His250 or His251 of the dimer-related FcRn heavy chain (Figure 2c). We previously showed that mutation of His250 and His251 results in a sixfold reduction in affinity for IgG (Raghavan et al., 1994). Two mutations were constructed (βE89H and βE89K) to test the role of Glu89 in dimer formation and IgG binding (Table 1). Similar to the mutations at Gly191, substitution of Glu89 with either a histidine or a lysine residue results in a slight reduction in high affinity IgG binding, and a reduction in the fraction of high affinity IgG binding sites (Figure 5c; Table 2). Unlike the Gly191 mutations, however, the total number of binding sites is not reduced, and the effect on affinity is largely independent of coupling density (Figure 5d; Table 2).
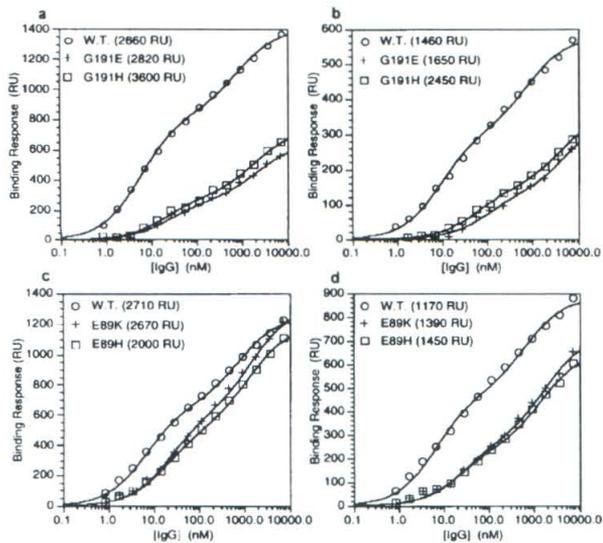
Figure 5. IgG binding to FcRn mutants at the dimer interface. Equilibrium IgG binding to wild-type and mutant FcRn proteins immobilized on a biosensor chip is plotted as a function of the log of the IgG concentration. The best fit binding curves (continuous lines), modelled as two classes of non-interacting binding sites, correspond to the values reported in Table 2. Experiments were performed at high (a and c) and low (b and d) coupling densities. Values for coupling densities are indicated in parentheses.

Discussion

Functionally critical residues for protein–protein interactions can be mapped by site-directed mutagenesis followed by determination of mutant binding affinities. This is an especially powerful method when high resolution structures are available to identify the residues at the receptor–ligand interface. One of the best characterized protein-protein interfaces is the interaction between human growth hormone (hGH) and its receptor (hGHR). Wells and co-workers have systematically changed all the residues that compose the crystallographically observed interface between the hormone and its receptor (De Vos *et al.*, 1992) to alanine residues, and find that only a small subset of the amino acids at the interface (approximately 5 of the 24 to 31 residues) are responsible for most of the binding affinity (Wells & De Vos, 1996). Thus there is a distinction between the "functional epitope" (those residues that exert a major effect upon the binding affinity; $\Delta\Delta G > 2$ kcal/mol for substitution of a single amino acid to alanine) and the "structural binding site" (all residues at the interface).

In the case of the interaction between FcRn and IgG, the structural binding site cannot be identified with certainty using the available low resolution FcRn/Fc co-crystal structure (Burmeister *et al.*, 1994b). The IgG binding site was tentatively identified as primarily formed from a surface of the α2 domain with additional potential contributions from residues in the α1 domain, the top surface of the β2m domain, and/or the α3 domain of a dimer related molecule (Burmeister *et al*, 1994b). Because the complex structure was determined using crystals that diffracted to only about 6.5 Å and a portion of the Fc C_H2 domains was disordered in the final electron density maps, we sought experimental confirmation of the IgG binding site on FcRn. The identification of several amino acid substitutions that substantially disrupt IgG binding within the structural epitope predicted from the low resolution co-crystal structure provides this confirmation. We have identified several amino acid mutations (W133A, E117S, D137N, E132Q&E135Q, βI1A and β-1Y) that result in a decrease in the free energy of binding of more than 2 kcal/mol (Table 2). These mutants, and all mutants in this study, are correctly folded as assayed by monoclonal antibody reactivity and heterodimer assembly (see Methods). Thus FcRn heavy chain residues Trp133, Glu117, Asp137, either Glu132, Glu135 or both, as well as Ile1 and the N terminus of β2m comprise much of the functional IgG binding epitope.

Well-characterized functional binding epitopes for protein–protein interfaces often include surface exposed hydrophobic residues that make important contributions to the free energy of binding. For example, hydrophobic interactions contribute to the binding energy in several antibody–antigen interactions (Dall'Acqua *et al.*, 1996; Kelley & O'Connell, 1993; Tsumoto *et al.*, 1995). In addition, the binding affinity for the hGH/hGHR interaction arises primarily from hydrophobic interactions, with electrostatic interactions being less important (Wells & De Vos, 1996). For hGHR, the largest contributions to the free energy of binding come from Trp104 and Trp169. Replacement of either residue with an alanine results in a loss of binding free energy of more than 4.5 kcal/mol. The next three largest contributions also are made by hydrophobic residues (Ile103, Ile105, and Pro106), each with a loss of binding free energy of 1.5 to 3.5 kcal/mol. For the hormone, both polar and hydrophobic residues show significant decreases in

binding free energy when substituted with alanine residues. Interestingly, several of the polar residues (e.g. Lys172, Thr175, and Arg43) use the aliphatic portions of their side-chains to contact the receptor.

There is also a strong hydrophobic component to the interaction between FcRn and its IgG ligand. Previous reports have implicated Ile253 as part of the ligand's functional epitope (Kim *et al.*, 1994a,b,c; Medesan *et al.*, 1997; Raghavan *et al.*, 1995a). Here we have shown that Trp133 is part of the receptor's functional epitope. From the crystal structure of the complex, these two hydrophobic side-chains are known to be positioned near each other (Figure 2a). Thus Ile253 of the ligand and Trp133 of the receptor are likely to form the hydrophobic core of a binding interface similar to those seen for hGH–hGHR and in several antibody–antigen interactions (Dall'Acqua *et al.*, 1996; Kelley & O'Connell, 1993; Tsumoto *et al.*, 1995; Wells & De Vos, 1996).

In addition to hydrophobic effects, the FcRn–Fc interface relies on strong electrostatic interactions to provide binding free energy. It was previously shown that His310 and one or more of histidine residues 433, 435, and 436 of the Fc ligand are necessary for full binding affinity and *in vivo* function (Kim *et al.*, 1994a,b,c; Medesan *et al.*, 1997; Raghavan *et al.*, 1994, 1995a). Here, we have identified three FcRn mutations, E117S, D137N, and E132Q&E135Q, that dramatically reduce binding affinity. Since these substitutions are conservative, most replacing only the negatively charged carboxylate group with a neutral amide group, it is likely that the disrupted interactions are salt bridges. In the FcRn/Fc cocrystal structure, each of the mutated acidic residues from FcRn are near one or more of the implicated Fc histidine residues (Figure 2b), suggesting that the protonated histidine residues form a pH-dependent salt bridge to the corresponding acidic residues on FcRn.

The approximately wild-type binding of the 84-86 mutant, suggests that these residues within the FcRn α1 domain do not contribute to the functional IgG binding site, while the substantially reduced binding of βI1A and β-1Y demonstrates that the β2m light chain does. The hydrophobic Ile side-chain at the N terminus of β2m is positioned near the side-chains of Fc residue 309 (Leu, Val, Gln, or Met in rat, murine, and human IgGs; (Kabat *et al.*, 1991)) and Fc residue 311 (Gln or Arg in rat, murine, and human IgGs; (Kabat *et al.*, 1991)), and could form a hydrophobic interaction with the aliphatic portions of one or the other of these side-chains. Extension of the N terminus by one residue by the addition of a bulky tyrosine residue (β-1Y) could reduce binding through a steric hindrance. Alternatively, the N-terminal extension could affect binding by blocking a specific interaction with the N-terminal amino group. In proteins, α-NH_2 groups have a pK_a that is generally ~8 (Fersht, 1985). In FcRn, the presumably protonated α-NH_2 group of β2m is positioned where it can form a hydrogen bond with the back-

bone carbonyl of residue 115 on the heavy chain and also a pH dependent salt-bridge with Glu117. Thus the protonated α-NH_2 group of β2m could help position Glu117 of the heavy chain to form an anionic binding site for His310 of the Fc ligand.

Inhibition studies using anti-β2m specific monoclonal antibodies previously suggested the involvement of the FcRn β2m domain in binding to IgG (Raghavan *et al.*, 1994). This suggestion is confirmed by the present demonstration that the N-terminal region of β2m constitutes a functional epitope for ligand binding. Crystal structures show that β2m interacts similarly with the heavy chains of class I MHC molecules and FcRn (Bjorkman & Parham, 1990; Burmeister *et al.*, 1994a). In both cases, the β2m and α3 immunoglobulin-like domains of FcRn and class I MHC molecules are related by a 152(± 7)° rotation followed by a ≈13 Å translation that positions the β2m domain underneath the α1-α2 domain platform (Burmeister *et al.*, 1994a and references therein). The conservation of this asymmetric α3–β2m interaction serves different functional purposes since the β2m position in FcRn is critical for interaction with its IgG ligand, yet the same β2m position is found in MHC molecules, which do not function as receptors for IgG.

The present study extends the results from previous work indicating that two FcRn molecules dimerize, as observed in FcRn and FcRn/Fc crystals (Burmeister *et al.*, 1994a,b), to form the high affinity binding site for a single IgG. We previously reported that FcRn mutants with alterations that specifically affect the interaction of IgG with FcRn dimers but not with monomers (residues 250 and 251 or the loop comprising residues 219 to 224; Figure 1) showed reduced affinities in biosensor assays and in measurements of labeled IgG binding to cell surface FcRn (Raghavan *et al.*, 1994). Here, we report two additional regions that affect FcRn dimerization and/or the interaction of Fc with the FcRn dimer. Mutations at FcRn residue 191, which interacts with its counterpart on the dimer related receptor, result in a lower affinity and fewer high affinity binding sites. These effects are enhanced at lower receptor density, suggesting that mutation of this residue exerts its effect through interference with FcRn dimer formation. By contrast, mutation of β2m residue Glu89 results in a slight reduction in binding affinity which is not dependent upon FcRn coupling density and is not accompanied by a significant reduction in the number of binding sites. These results suggest that β2m Glu89 interacts primarily with the IgG ligand directly rather than the dimer-related FcRn heavy chain. Disorder of the Fc C_H2 domain makes it difficult to identify potential contacts for Glu89 from the co-crystal structure (Burmeister *et al.*, 1994b). However, a direct interaction between the Fc C_H2 domain and residues at the dimer interface near β2m Glu89 (FcRn His250 and His251) was suggested by modeling studies using a computational docking algorithm to predict the orien-

tation of the Fc C_H2 domain when bound to FcRn (Weng *et al.*, 1997). If Fc becomes distorted upon binding to the FcRn dimer as predicted by the modeling studies, β2m Glu89 could contact the C_H2 domain.

Conclusions

We have designed and constructed several amino acid substitutions on FcRn to map its binding site for IgG. The identification of several mutations that reduce the IgG affinity confirm and further define the IgG binding site on FcRn that was identified at low resolution in the FcRn–Fc co-crystal structure (Burmeister *et al.*, 1994b). These functionally important IgG binding epitopes on FcRn complement previously reported binding epitopes on Fc (Kim *et al.*, 1994a,b,c; Medesan *et al.*, 1997; Raghavan *et al.*, 1994, 1995a). Ile253 of Fc and Trp133 of FcRn are likely to form the core of a strong hydrophobic interaction, while Glu117, Glu132, Glu135 and Asp137 on the receptor provide anionic binding partners for the protonated forms of His310, His433, His435 and His436 of the ligand. Finally, amino acid substitutions for Gly191 at the dimer interface display an FcRn concentration dependent effect on IgG affinity, providing additional evidence that FcRn dimerization is necessary for normal high affinity IgG binding.

Methods

Reagents

1B5, a mouse IgG1 monoclonal antibody against human Zn-α2-glycoprotein (Sánchez *et al.*, 1997) whose interaction with wild-type FcRn has been previously characterized (Vaughn & Bjorkman, 1997), was used for biosensor binding assays. Anti-CD4 (rat IgG2a) used in the cell binding assays was purchased from Boehringer Mannheim. Its interaction with wild-type FcRn has also been characterized in a biosensor assay (Vaughn & Bjorkman, 1997). Two mouse IgG1 anti-FcRn monoclonal antibodies, 1G3 and 4C11 (Raghavan *et al.*, 1994), were used for immunoaffinity chromatography, enzyme-linked immunosorbent assays (ELISAs), immunoprecipitations, and fluorescence-activated cell sorting (FACS). Rabbit anti-human β2m and peroxidase-conjugated anti-mouse, as well as goat fluorescein isothiocyanate (FITC)-conjugated anti-rabbit antibodies were purchased from Boehringer Mannheim. Iodobeads were obtained from Pierce. ^{125}I (specific activity ≈100 mCi/ml) was obtained from Amersham in the form of sodium iodide. *N*-ethyl-*N*′-(3-diethylamino-propyl)-carbodiimide (EDC), *N*-hydroxysuccinimide (NHS), BIAcore surfactant P20, and CM5 sensor chips were obtained from BIAcore AB. CNBr-activated Sepharose and PD-10 columns were purchased from Pharmacia.

Construction and expression of mutant FcRn molecules

Locations for FcRn mutants were chosen using the crystal structures of FcRn alone (Protein Data Bank entry 1FRU) and a complex between FcRn and Fc (Protein Data Bank entry 1FRT). For expression of soluble mutant FcRn heterodimers, the cDNAs encoding the truncated FcRn heavy chain (residues 1 to 269) (Gastinel *et al.*, 1992) or rat β2m were altered by site-directed mutagenesis (Kunkel *et al.*, 1987). After verifying the sequence of the altered cDNAs, the mutant cDNAs were subcloned into the expression vector pBJ5-GS which carries the glutamine synthetase gene as a means of selection and amplification in the presence of the drug methionine sulfoximine (Bebbington & Hentschel, 1987). Co-transfection of FcRn and β2m expression vectors into Chinese hamster ovary (CHO) cells, selection, and amplification were carried out as described (Gastinel *et al.*, 1992; Raghavan *et al.*, 1994). Cell supernatants were screened for FcRn heterodimer expression by a sandwich ELISA, using either the FcRn heavy chain-specific monoclonal antibody 1G3 or 4C11 as the capture antibody and a polyclonal antiserum against human β2m to detect positive samples as described (Raghavan *et al.*, 1995b), and confirmed by immunoprecipitation with the other antibody. Mutant proteins that retained functional binding to IgG were purified from the supernatants of the highest expressing clones by pH dependent binding to a rat IgG column (Gastinel *et al.*, 1992). Mutants that did not bind IgG sufficiently for IgG-affinity chromatography were purified using an immunoaffinity column constructed with an anti-FcRn monoclonal antibody. Purified 1G3 antibody was covalently attached to CNBr-activated Sepharose following the manufacturer's protocol. Binding and elution conditions for FcRn binding to the 1G3 column were identified using a biosensor assay of wild type FcRn interacting with amine coupled 1G3 at pH 8 (see below). Supernatants from cells expressing wild-type or mutant FcRn proteins were passed over the 1G3 column at pH 7.5, which was subsequently washed with ≈ten column volumes of 50 mM sodium phosphate (pH 7.5 to 8.5) and eluted with 50 mM sodium phosphate (pH 3.0). Eluates were immediately neutralized with 1 M disodium phosphate. These elution conditions did not affect the ability of wild-type FcRn to bond again to the 1G3 column or the behavior of wild-type FcRn in subsequent biosensor assays (data not shown).

For expression of membrane-bound FcRn mutants, mutations were introduced into a previously described lipid-linked version of rat β2m that pairs with the truncated FcRn heavy chain at the surface of transfected cells (Gastinel *et al.*, 1992). The lipid-linked β2m protein consists of the phosphatidylinositol anchoring signal of decay accelerating factor (DAF; residues 311 to 347; (Caras *et al.*, 1987)) fused to the final amino acid of

β2m (β2m-DAF). Altered β2m-DAF cDNAs were sequenced and subcloned into the PBJ5-GS expression vector, which was co-transfected into CHO cells along with the expression vector encoding wild type soluble FcRn. After two to four weeks, populations were screened for heterodimer expression by immunostaining at pH 8.0 with the anti-FcRn antibody 1G3 and FITC-conjugated goat anti-mouse secondary antibody. High expressing lines were isolated by single cell FACS. Cell sorting and analysis was performed on a Coulter Epics Elite flow cytometer using Coulter Elite software version 3.0.

The correct folding of each of the FcRn mutants used in this study was demonstrated by several experiments. First, each of the mutants is either expressed on the cell surface or secreted, therefore none of the mutations result in intracellular retention as would be expected for misfolded proteins. Secondly, all mutants assemble into heterodimers, as verified either by SDS-PAGE analysis of purified protein and in a sandwich ELISA for secreted mutants, or by cell surface reactivity with an anti-FcRn heavy chain monoclonal antibody for the mutations introduced into lipid-linked β2m and expressed as heterodimers with soluble FcRn heavy chains (data not shown). Finally, each mutant retains the immunoreactivity of wild-type FcRn for monoclonal antibodies generated against correctly folded wild-type FcRn (4C11 and 1G3; Raghavan *et al.*, 1994). All secreted mutants were selected using a sandwich ELISA and immunoprecipitation employing 4C11 and 1G3, non-IgG binding mutants were purified on a 1G3 immunoaffinity column, and cells expressing lipid-linked mutants were isolated by flow cytometry using 1G3. The correct folding of each mutant was expected since each of the mutations was introduced into residues that are solvent exposed in the crystal structure of FcRn (Burmeister *et al.*, 1994a).

Purification and iodination of IgG

Purified anti CD4 was iodinated using iodo-beads according to the manufacturer's protocol. Briefly, one or two beads were rinsed in 1 ml of 100 mM sodium phosphate buffer (pH 6.0), dried on filter paper and added to a mixture of 20 μl of ^{125}I in 180 μl of the same buffer. After five minutes, 0.5 ml (0.4 mg) of purified IgG was added and the mixture was allowed to incubate for approximately 15 minutes at room temperature. Unreacted ^{125}I was removed by gel filtration using a PD-10 column.

Cell binding assays

Cells expressing wild-type FcRn/β2m-DAF or mutant FcRn/β2m-DAF were grown to confluence in tissue culture plates and assayed for IgG binding as described (Raghavan *et al.*, 1994). Cells were detached, pelleted, washed and resuspended in binding buffer (Hank's balanced salt solution,

10 mM Hepes, 0.25% bovine serum albumin, pH 6.0) to a concentration of approximately 3×10^6 cells/ml. In triplicate assays 3×10^5 to 5×10^5 cells were mixed with various amounts of ^{125}I labeled IgG and binding buffer to a total volume of 0.5 ml. The samples were incubated for at least two hours at room temperature, then pelleted and the supernatants removed and set aside. Cell pellets were washed with 1.0 ml of cold binding buffer. The levels of radioactivity in the supernatants and cell pellets were measured using a Beckman gamma 5500 counter. Non-specific binding was determined by similar treatment of untransfected CHO cells. The concentration of free IgG was determined from the radioactivity in the supernatant and the specific activity of the labeled IgG. The binding response was similarly calculated from the radioactivity in the washed cell pellet.

Biosensor experiments

A BIAcore biosensor system (Biacore AB) was used to evaluate the interaction of secreted mutant and wild type FcRn proteins with IgG as previously reported (Vaughn & Bjorkman, 1997; and references therein). Briefly, wild-type or mutant FcRn proteins were immobilized to a biosensor chip surface using standard amine coupling chemistry as described in the BIAcore manual, and different concentrations of purified IgG were injected over the immobilized FcRn. IgG concentrations were determined spectrophotometrically using an extinction coefficient at 280 nm of $216,000 \ M^{-1} \ cm^{-1}$ (IgG; Fasman, 1989). Sensor chip surfaces were regenerated by injecting a pulse of 50 mM phosphate, 150 mM NaCl (pH 8). Variation in coupling densities was achieved by varying the chip activation step from four to seven minutes.

Because systematic errors can contribute to the uncertainty of K_D determinants (including slight variations in the pH of the binding buffer, the specifics of the coupling reaction, the concentration of the ligand, and the specific batch and purification state of the ligand), K_D values can be more accurately compared for experiments conducted in parallel on the same chip. For each experiment, wild-type FcRn, one or two mutants, and buffer without protein were coupled in one of the four available flowcells. The same dilution series of ligand and a buffer blank were passed over each flow cell, and the response at ten minutes (taken as the equilibrium response) was recorded after subtraction of the buffer response. A modest response was observed for higher ligand concentrations ($> \approx 62$ nM) in mock-coupled flow cells, reaching a maximum of ≈200 RU at the highest ligand concentration of 8 μM. This blank response, most likely due to either the bulk refractive index contribution of the ligand and/or a nonspecific low affinity interaction with the dextran matrix of the chip, was subtracted from the equilibrium responses recorded for FcRn coupled flow cells to yield a net equilibrium response.

Calculation of binding affinities and changes in the free energy of binding

For both the cell binding and biosensor assays, the equilibrium binding response was plotted as a function of the log of free IgG concentration. These data were fit to either a model of one or two non-interacting binding sites by non-linear regression as previously described (Vaughn & Bjorkman, 1997). In each case in which a model of two non-interacting sites is used, the two site model produces a lower cross-validated residual (Vaughn & Bjorkman, 1997) than the one site model. K_D and R_{max} (response corresponding to complete binding) values are presented when a single class of non-interacting sites was modeled, and $K_{D,1}$ (high affinity binding constant), $K_{D,2}$ (low-affinity binding constant), f_1 (fraction of total binding sites corresponding to the high affinity class of binding sites) and $R_{max,tot}$ (response corresponding to complete occupancy of both classes of binding sites) values are presented when two classes of non-interacting sites were modeled (Table 2). For each of these parameters, standard errors are also reported as determined in the KaleidaGraph® implementation of the Levenberg–Marquardt algorithm, which reflect the precision of individual curve fittings (Table 2). The reproducibility of the derived affinities is $\pm \approx 20\%$ as evaluated by comparing the six independent K_D values reported for wild-type FcRn from experiments conducted on different biosensor chips (Table 2). $\Delta\Delta G$ values were calculated as $2.303RT \log(K_{mut}/K_{w.t})$ for the high affinity binding constant $(K_{D,1})$ where R is the gas constant $(1.99 \times 10^{-3} \text{ kcal mol}^{-1} \text{ K}^{-1})$ and T is the absolute temperature (295 K). Results from biosensor experiments in Table 2 correspond to the interaction of wild type and mutant FcRn proteins with the 1B5 monoclonal IgG. Similar results were obtained for experiments on FcRn mutants that interact directly with IgG using the anti-CD4 IgG (data not shown).

References

Bebbington, C. R. & Hentschel, C. C. G. (1987). The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells. In *DNA Cloning: A Practical Approach* (Glover, D. M., ed.), pp. 163–188, IRL Press, Oxford.

Bjorkman, P. J. & Parham, P. (1990). Structure, function and diversity of class I major histocompatibility complex molecules. *Annu. Rev. Biochem.* **90**, 253–288.

Burmeister, W. P., Gastinel, L. N., Simister, N. E., Blum, M. L. & Bjorkman, P. J. (1994a). Crystal structure at 2.2 Å resolution of the MHC-related neonatal Fc receptor. *Nature*, **372**, 336–343.

Burmeister, W. P., Huber, A. H. & Bjorkman, P. J. (1994b). Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature*, **372**, 379–383.

Caras, I. W., Weddell, G. N., Davitz, M. A., Nussenzweig, V. & Martin, D. W. (1987). Signal for attachment of a phospholipid membrane anchor in decay accelerating factor. *Science*, **238**, 1280–1283.

Dall'Acqua, W., Goldman, E. R., Eisenstein, E. & Mariuzza, R. A. (1996). A mutational analysis of the binding of two different proteins to the same antibody. *Biochemistry*, **35**, 9667–9676.

De Vos, A. M., Ultsch, M. & Kossiakoff, A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, **255**, 306–312.

Fasman, G. D. (1989). Editor of *Practical Handbook of Biochemistry and Molecular Biology*, p. 265, CRC Press, Boca Raton, FL.

Fersht, A. (1985). *Enzyme Structure and Mechanism*, 2nd edit., W. H. Freeman and Company, New York, USA.

Gastinel, L. N., Simister, N. E. & Bjorkman, P. J. (1992). Expression and crystallization of a soluble and functional form of an Fc receptor related to class I histocompatibility molecules. *Proc. Natl Acad. Sci. USA*, **89**, 638–642.

Junghans, R. P. (1991). Finally! The Brambell Receptor (FcRB). Mediator of transmission of immunity and protection from catabolism for IgG. *Immunol. Res.* **16**, 29–57.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). Sequences of proteins of immunological interest. US Department of Health and Human Services, Bethesda, MD, USA.

Kelley, R. F. & O'Connell, M. P. (1993). Thermodynamic analysis of an antibody functional epitope. *Biochemistry*, **32**, 6828–6835.

Kim, J.-K., Tsen, M.-F., Ghetie, V. & Ward, E. S. (1994a). Catabolism of the murine IgG1 molecule: evidence that both CH2-CH3 domain interfaces are required for the persistence of IgG1 in the circulation of mice. *Scand. J. Immunol.* **40**, 457–465.

Kim, J.-K., Tsen, M.-F., Ghetie, V. & Ward, E. S. (1994b). Localization of the site of the murine IgG1 molecule that is involved in binding to the murine intestinal Fc receptor. *Eur. J. Immunol.* **24**, 2429–2434.

Kim, J.-K., Tsen, M.-F., Ghetie, V. & Ward, E. S. (1994c). Identifying amino acid residues that influence plasma clearance of mouse IgG1 fragments by site directed mutagenesis. *Eur. J. Immunol.* **24**, 542–548.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site directed mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367–382.

Mackenzie, N. (1984). Fc receptor-mediated transport of immunoglobulin across the intestinal epithelium of the neonatal rodent. *Immunol. Today*, **5**, 364–366.

Medesan, C., Matesoi, D., Radu, C., Ghetie, V. & Ward, E. S. (1997). Delineation of the amino acid residues involved in transcytosis and catabolism of mouse IgG. *J. Immunol.* **158**, 2211–2217.

Raghavan, M. & Bjorkman, P. J. (1996). Fc receptors and their interactions with immunoglobulins. *Annu. Rev. Cell Biol.* **12**, 181–220.

Raghavan, M., Chen, M. Y., Gastinel, L. N. & Bjorkman, P. J. (1994). Identification of interaction sites in the class I NHC-related Fc receptor/immunoglobulin G complex. *Immunity*, **1**, 303–315.

Raghavan, M., Bonagura, V. R., Morrison, S. L. & Bjorkman, P. J. (1995a). Analysis of the pH dependence of the neonatal Fc receptor/immunoglobulin G interaction using antibody and receptor variants. *Biochemistry*, **34**, 14649–14657.

Raghavan, M., Wang, Y. & Bjorkman, P. J. (1995b). Effects of receptor dimerization on the interaction between class I MHC related Fc receptor and immunoglobulin G. *Proc. Natl Acad. Sci. USA*, **92**, 11200–11204.

Rodewald, R. (1976). pH-dependent binding of immunoglobulins to intestinal cells of the neonatal rat. *J. Cell Biol.* **71**, 666–670.

Rodewald, R. & Kraehenbuhl, J.-P. (1984). Receptor-mediated transport of IgG. *J. Cell Biol.* **99**, S159–S164.

Sánchez, L. M., López-Otín, C. & Bjorkman, P. J. (1997). Characterization and crystallization of human Zn-α2-glycoprotein, a soluble class I MHC homolog. *Proc. Natl Acad. Sci. USA*, **94**, 4626–4630.

Simister, N. E. & Mostov, K. E. (1989). An Fc receptor structurally related to MHC class I antigens. *Nature*, **337**, 184–187.

Simister, N. E. & Rees, A. R. (1985). Isolation and characterization of an Fc receptor from neonatal rat small intestine. *Eur. J. Immunol.* **15**, 733–738.

Tsumoto, K., Ogasahara, K., Ueda, Y., Watanabe, K., Yutani, K. & Kumagai, I. (1995). Role of Tyr residues in the contact region of anti-lysozyme monoclonal antibody HyHEL10 for antigen binding. *J. Biol. Chem.* **270**, 18551–18557.

Vaughn, D. E. & Bjorkman, P. J. (1997). High affinity binding of the neonatal Fc receptor to its IgG ligand requires receptor immobilization. *Biochemistry*, **36**, 9374–9380.

Wells, J. A. & De Vos, A. M. (1996). Hematopoietic receptor complexes. *Annu. Rev. Biochem.* **65**, 609–634.

Appendix B:

Protein Interface Properties

This appendix contains the data table of the quantifiable aspects of the 171

nonobligate protein-protein interfaces used for the analysis made in Chapter 5.

	pdbid			Resolution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Protease-protease inhibitor								
Chymotrypsin	1acb	A	E	3	25198	43430	10380	7.3E+02
elgin C	1acb	B	I	3	7315	13630	4189	8.1E+02
Thrombin	1avg	A	H L	1.2	34453	59510	13600	7.2E+02
triabin	1avg	B	I	1.2	15984	28400	7474	6.5E+02
Trypsin	1avw	A	A	3.2	23280	39400	9084	8.1E+02
soybean trypsin inhibitor	1avw	B	B	3.2	17958	32210	8361	9.2E+02
Collagenase	1azz	A	A B	2.8	46926	79930	18520	1.4E+03
ecotin	1azz	B	C	2.8	14868	29020	9283	1.5E+03
Trypsin var. D189G, G226D	1brc	A	E	3	23418	40400	9375	5.9E+02
amyloid beta-protein precursor	1brc	B	I	3	5909	10950	3476	7.2E+02
Thrombin	1bth	A	L H	3.5	32674	56350	13330	1.1E+03
bovine pancreatic trypsin inhibitor	1bth	B	P	3.5	6510	11990	3753	1.2E+03
Trypsin	1c9t	A	A	3.5	23269	39350	9076	5.3E+02
Bdellstasin	1c9t	B	G	3.5	5627	11000	3896	7.0E+02
Chymotrypsin	1ca0	A	F G H	2.9	10052	19900	6921	4.3E+02
amyloid beta-protein precursor	1ca0	B	I	2.9	5988	10980	3586	4.6E+02
Chymotrypsin	1cbw	A	F G H	2.3	25028	42980	10050	6.4E+02
Bovine pancreatic trypsin inhibitor	1cbw	B	I	2.3	6510	12240	3926	7.5E+02
Chymotrypsin	1cgi	A	E	1.9	25625	44180	10470	9.4E+02
trypsin inhibitor var 3	1cgi	B	I	1.9	6327	12040	3995	1.1E+03
Chymotrypsin	1cgj	A	E	1.7	25625	44130	10470	9.3E+02
trypsin inhibitor var 4	1cgj	B	I	1.7	6277	11860	3920	1.0E+03
Chymotrypsin	1cho	A	E	3	24925	42820	9920	6.6E+02
turkey ovomucoid	1cho	B	I	3	5753	10880	3629	7.9E+02
Subtilisin Carlsberg	1cse	A	E	2.4	27246	45260	9696	6.4E+02
elgin C	1cse	B	I	2.4	7315	13740	4125	8.3E+02
Factor VIIa	1dan	A	L H	2.4	42615	76800	20470	1.9E+03
soluble tissue factor	1dan	B	T U	2.4	21986	40010	10790	1.8E+03
Elastase	1eai	A	B	2.9	25877	46470	11010	9.5E+02
Chymotrypsin/elastase inhibitor	1eai	B	D	2.9	6583	12890	4526	1.1E+03
ecotin	1ezs	A	A B	1.65	27136	51820	14960	1.2E+03
Trypsin II anionic	1ezs	B	C	1.65	23753	40760	9408	1.1E+03
anti-trypsin	1ezx	A	A	2.7	41148	70790	15090	4.5E+02
Trypsin	1ezx	B	B	2.7	14246	25930	7011	3.9E+02
Natural Killer Cell protease	1fi8	A	A	1.9	23738	41600	10130	7.3E+02
ecotin	1fi8	B	C D	1.9	27325	27110	8617	7.7E+02
Elastase	1fle	A	E	3.7	25877	44440	10490	8.1E+02
elafin	1fle	B	I	3.7	5005	9882	3548	9.5E+02
Caspase activator	1g73	A	A B	2.5	34806	64290	19120	9.8E+02
inhibitor	1g73	B	C	2.5	10277	18340	5219	8.3E+02
Kallikrein	1hia	A	A B	3.25	25365	43280	10060	7.7E+02
hirustasin	1hia	B	I	3.25	5180	10060	3608	9.6E+02
Factor Xa	1kig	A	H L	2.9	32377	55090	13320	8.5E+02
anti-coagulant	1kig	B	I	2.9	6953	12830	4214	9.2E+02
Trypsin	1mct	A	A	2.3	23440	39690	9025	6.7E+02
bitter gourd inhibitor	1mct	B	I	2.3	3153	6256	2323	8.3E+02
Mesenterocopeptidase	1mee	A	A	2.6	27613	45960	9819	8.1E+02
elgin C	1mee	B	I	2.6	7428	13850	4091	9.4E+02
Thrombin	1mkw	A	L H	2.5	32728	55920	12730	6.1E+02
prethrombin-2	1mkw	B	K	2.5	32983	57630	14190	6.6E+02

	pdbid			Circularity	RMSD to plane	Atom burial	Atom accessibility
Protease-protease inhibitor							
Chymotrypsin	1acb	A	E	6.3E-01	2.9E+00	7.3E+00	1.2E+01
elgin C	1acb	B	I	5.1E-01	2.1E+00	1.0E+01	1.6E+01
Thrombin	1avg	A	H L	3.5E-01	2.2E+00	9.5E+00	1.7E+01
triabin	1avg	B	I	3.3E-01	2.2E+00	7.7E+00	1.4E+01
Trypsin	1avw	A	A	6.4E-01	2.8E+00	7.0E+00	9.6E+00
soybean trypsin inhibitor	1avw	B	B	5.2E-01	2.7E+00	1.1E+01	1.4E+01
Collagenase	1azz	A	A B	8.4E-02	3.8E+00	7.7E+00	1.2E+01
ecotin	1azz	B	C	6.0E-02	3.2E+00	1.1E+01	1.5E+01
Trypsin var. D189G, G226D	1brc	A	E	6.3E-01	2.8E+00	6.6E+00	9.8E+00
amyloid beta-protein precursor	1brc	B	I	3.7E-01	2.1E+00	1.2E+01	1.7E+01
Thrombin	1bth	A	L H	7.7E-01	4.2E+00	7.5E+00	1.1E+01
bovine pancreatic trypsin inhibitor	1bth	B	P	7.0E-01	3.5E+00	1.1E+01	1.4E+01
Trypsin	1c9t	A	A	3.7E-01	2.7E+00	6.1E+00	9.4E+00
Bdellstasin	1c9t	B	G	1.9E-01	2.2E+00	1.3E+01	1.8E+01
Chymotrypsin	1ca0	A	F G H	5.0E-01	2.3E+00	7.9E+00	1.3E+01
amyloid beta-protein precursor	1ca0	B	I	5.0E-01	2.1E+00	9.8E+00	1.5E+01
Chymotrypsin	1cbw	A	F G H	5.3E-01	2.9E+00	7.3E+00	1.2E+01
Bovine pancreatic trypsin inhibitor	1cbw	B	I	2.6E-01	2.5E+00	1.1E+01	1.5E+01
Chymotrypsin	1cgi	A	E	6.3E-01	3.1E+00	7.6E+00	1.1E+01
trypsin inhibitor var 3	1cgi	B	I	3.2E-01	2.9E+00	1.1E+01	1.5E+01
Chymotrypsin	1cgj	A	E	6.0E-01	3.0E+00	7.5E+00	1.2E+01
trypsin inhibitor var 4	1cgj	B	I	3.0E-01	2.7E+00	1.1E+01	1.5E+01
Chymotrypsin	1cho	A	E	4.8E-01	2.7E+00	7.1E+00	1.1E+01
turkey ovomucoid	1cho	B	I	2.9E-01	2.2E+00	1.1E+01	1.6E+01
Subtilisin Carlsberg	1cse	A	E	5.9E-01	2.6E+00	7.0E+00	1.2E+01
elgin C	1cse	B	I	4.6E-01	3.0E+00	1.2E+01	1.7E+01
Factor VIIa	1dan	A	L H	1.7E-01	4.3E+00	1.1E+01	1.7E+01
soluble tissue factor	1dan	B	T U	1.5E-01	5.1E+00	8.7E+00	1.6E+01
Elastase	1eai	A	B	5.3E-01	3.6E+00	8.0E+00	1.3E+01
Chymotrypsin/elastase inhibitor	1eai	B	D	5.2E-01	2.9E+00	1.1E+01	1.7E+01
ecotin	1ezs	A	A B	6.6E-01	3.5E+00	1.1E+01	1.7E+01
Trypsin II anionic	1ezs	B	C	8.2E-01	3.5E+00	7.2E+00	1.2E+01
anti-trypsin	1ezx	A	A	4.5E-01	1.4E+00	1.1E+01	1.7E+01
Trypsin	1ezx	B	B	7.5E-01	1.6E+00	8.6E+00	1.6E+01
Natural Killer Cell protease	1fi8	A	A	5.3E-01	3.0E+00	7.8E+00	1.2E+01
ecotin	1fi8	B	C D	4.1E-01	2.9E+00	1.0E+01	1.7E+01
Elastase	1fle	A	E	8.4E-01	3.4E+00	7.8E+00	1.3E+01
elafin	1fle	B	I	6.9E-01	2.7E+00	1.1E+01	1.8E+01
Caspase activator	1g73	A	A B	3.9E-01	3.7E+00	1.1E+01	1.8E+01
inhibitor	1g73	B	C	5.2E-01	3.9E+00	7.2E+00	1.5E+01
Kallikrein	1hia	A	A B	4.1E-01	3.1E+00	7.8E+00	1.1E+01
hirustasin	1hia	B	I	2.3E-01	3.1E+00	1.2E+01	1.8E+01
Factor Xa	1kig	A	H L	4.5E-01	3.8E+00	7.4E+00	1.1E+01
anti-coagulant	1kig	B	I	5.3E-01	3.6E+00	1.1E+01	1.4E+01
Trypsin	1mct	A	A	5.4E-01	2.8E+00	6.2E+00	9.4E+00
bitter gourd inhibitor	1mct	B	I	2.3E-01	2.9E+00	1.2E+01	1.6E+01
Mesenterocopeptidase	1mee	A	A	6.2E-01	2.7E+00	7.6E+00	1.1E+01
elgin C	1mee	B	I	4.1E-01	2.9E+00	1.0E+01	1.4E+01
Thrombin	1mkw	A	L H	6.9E-01	2.3E+00	6.8E+00	1.3E+01
prethrombin-2	1mkw	B	K	6.9E-01	2.2E+00	9.1E+00	1.9E+01

Protease-protein inhibitor	pdbid			Shape complementarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Chymotrypsin	1acb	A	E	7.0E-01	1.0E+00	1.9E+00
elgin C	1acb	B	I	7.0E-01	1.0E+00	1.8E+00
Thrombin	1avg	A	H L	5.6E-01	1.0E+00	2.2E+00
triabin	1avg	B	I	5.6E-01	1.1E+00	2.1E+00
Trypsin	1avw	A	A	7.6E-01	1.0E+00	1.6E+00
soybean trypsin inhibitor	1avw	B	B	7.6E-01	1.2E+00	1.7E+00
Collagenase	1azz	A	A B	7.6E-01	9.9E-01	1.8E+00
ecotin	1azz	B	C	7.6E-01	9.7E-01	2.1E+00
Trypsin var. D189G, G226D	1brc	A	E	7.1E-01	1.0E+00	1.6E+00
amyloid beta-protein precursor	1brc	B	I	7.1E-01	1.1E+00	2.0E+00
Thrombin	1bth	A	L H	6.8E-01	1.0E+00	1.9E+00
bovine pancreatic trypsin inhibitor	1bth	B	P	6.8E-01	1.2E+00	2.1E+00
Trypsin	1c9t	A	A	7.2E-01	9.2E-01	2.0E+00
Bdellastasin	1c9t	B	G	7.2E-01	1.2E+00	2.3E+00
Chymotrypsin	1ca0	A	F G H	7.4E-01	8.7E-01	1.6E+00
amyloid beta-protein precursor	1ca0	B	I	7.4E-01	9.2E-01	1.9E+00
Chymotrypsin	1cbw	A	F G H	6.9E-01	1.1E+00	1.9E+00
Bovine pancreatic trypsin inhibitor	1cbw	B	I	6.9E-01	1.1E+00	1.9E+00
Chymotrypsin	1cgi	A	E	7.2E-01	1.2E+00	1.8E+00
trypsin inhibitor var 3	1cgi	B	I	7.2E-01	1.1E+00	1.9E+00
Chymotrypsin	1cgj	A	E	7.0E-01	1.2E+00	1.8E+00
trypsin inhibitor var 4	1cgj	B	I	7.0E-01	1.1E+00	2.1E+00
Chymotrypsin	1cho	A	E	7.1E-01	1.0E+00	1.7E+00
turkey ovomucoid	1cho	B	I	7.1E-01	1.1E+00	1.7E+00
Subtilisin Carlsberg	1cse	A	E	7.3E-01	1.1E+00	1.6E+00
elgin C	1cse	B	I	7.3E-01	1.2E+00	1.7E+00
Factor VIIa	1dan	A	L H	6.6E-01	1.0E+00	1.9E+00
soluble tissue factor	1dan	B	T U	6.6E-01	1.1E+00	1.7E+00
Elastase	1eai	A	B	6.8E-01	1.1E+00	2.2E+00
Chymotrypsin/elastase inhibitor	1eai	B	D	6.8E-01	1.1E+00	2.5E+00
ecotin	1ezs	A	A B	7.3E-01	8.8E-01	2.2E+00
Trypsin II anionic	1ezs	B	C	7.3E-01	9.6E-01	1.8E+00
anti-trypsin	1ezx	A	A	5.7E-01	1.1E+00	1.8E+00
Trypsin	1ezx	B	B	5.7E-01	1.1E+00	2.1E+00
Natural Killer Cell protease	1fi8	A	A	6.8E-01	1.0E+00	2.1E+00
ecotin	1fi8	B	C D	6.8E-01	1.2E+00	2.1E+00
Elastase	1fle	A	E	7.1E-01	1.0E+00	2.0E+00
elafin	1fle	B	I	7.1E-01	1.2E+00	1.9E+00
Caspase activator	1g73	A	A B	7.6E-01	1.0E+00	1.8E+00
inhibitor	1g73	B	C	7.6E-01	1.1E+00	2.2E+00
Kallikrein	1hia	A	A B	7.3E-01	1.1E+00	1.8E+00
hirustasin	1hia	B	I	7.3E-01	1.1E+00	2.2E+00
Factor Xa	1kig	A	H L	7.9E-01	1.0E+00	1.8E+00
anti-coagulant	1kig	B	I	7.9E-01	1.3E+00	2.1E+00
Trypsin	1mct	A	A	7.8E-01	9.9E-01	1.6E+00
bitter gourd inhibitor	1mct	B	I	7.8E-01	1.1E+00	2.0E+00
Mesenterocopeptidase	1mee	A	A	7.7E-01	1.2E+00	1.5E+00
elgin C	1mee	B	I	7.7E-01	1.1E+00	1.6E+00
Thrombin	1mkw	A	L H	6.9E-01	1.1E+00	1.8E+00
prethrombin-2	1mkw	B	K	6.9E-01	7.1E-01	1.7E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
Protease-protease inhibitor							
Chymotrypsin	1acb	A	E	1.1E+01	1.5E+00	2.1E+01	2.9E+00
elgin C	1acb	B	I	1.1E+01	1.4E+00	2.1E+01	2.6E+00
Thrombin	1avg	A	H L	1.1E+01	1.5E+00	2.4E+01	3.3E+00
triabin	1avg	B	I	1.1E+01	1.7E+00	2.4E+01	3.7E+00
Trypsin	1avw	A	A	1.0E+01	1.2E+00	2.2E+01	2.7E+00
soybean trypsin inhibitor	1avw	B	B	1.0E+01	1.1E+00	2.2E+01	2.4E+00
Collagenase	1azz	A	A B	8.0E+00	5.9E-01	1.8E+01	1.3E+00
ecotin	1azz	B	C	8.0E+00	5.3E-01	1.8E+01	1.2E+00
Trypsin var. D189G, G226D	1brc	A	E	1.1E+01	1.9E+00	1.6E+01	2.7E+00
amyloid beta-protein precursor	1brc	B	I	1.1E+01	1.5E+00	1.6E+01	2.2E+00
Thrombin	1bth	A	L H	1.0E+00	8.9E-02	1.0E+00	8.9E-02
bovine pancreatic trypsin inhibitor	1bth	B	P	1.0E+00	8.1E-02	1.0E+00	8.1E-02
Trypsin	1c9t	A	A	1.2E+01	2.2E+00	1.9E+01	3.6E+00
Bdellstasin	1c9t	B	G	1.2E+01	1.7E+00	1.9E+01	2.7E+00
Chymotrypsin	1ca0	A	F G H	2.2E+01	5.1E+00	3.1E+01	7.2E+00
amyloid beta-protein precursor	1ca0	B	I	2.2E+01	4.8E+00	3.1E+01	6.7E+00
Chymotrypsin	1cbw	A	F G H	2.2E+01	3.5E+00	4.0E+01	6.3E+00
Bovine pancreatic trypsin inhibitor	1cbw	B	I	2.2E+01	2.9E+00	4.0E+01	5.4E+00
Chymotrypsin	1cgi	A	E	9.0E+00	9.6E-01	1.9E+01	2.0E+00
trypsin inhibitor var 3	1cgi	B	I	9.0E+00	8.3E-01	1.9E+01	1.7E+00
Chymotrypsin	1cgj	A	E	6.0E+00	6.5E-01	1.4E+01	1.5E+00
trypsin inhibitor var 4	1cgj	B	I	6.0E+00	5.7E-01	1.4E+01	1.3E+00
Chymotrypsin	1cho	A	E	1.1E+01	1.7E+00	1.9E+01	2.9E+00
turkey ovomucoid	1cho	B	I	1.1E+01	1.4E+00	1.9E+01	2.4E+00
Subtilisin Carlsberg	1cse	A	E	2.0E+01	3.1E+00	3.5E+01	5.5E+00
elgin C	1cse	B	I	2.0E+01	2.4E+00	3.5E+01	4.2E+00
Factor VIIa	1dan	A	L H	1.0E+01	5.3E-01	1.9E+01	1.0E+00
soluble tissue factor	1dan	B	T U	1.0E+01	5.6E-01	1.9E+01	1.1E+00
Elastase	1eai	A	B	7.0E+00	7.4E-01	1.5E+01	1.6E+00
Chymotrypsin/elastase inhibitor	1eai	B	D	7.0E+00	6.4E-01	1.5E+01	1.4E+00
ecotin	1ezs	A	A B	1.3E+01	1.0E+00	2.6E+01	2.1E+00
Trypsin II anionic	1ezs	B	C	1.3E+01	1.2E+00	2.6E+01	2.3E+00
anti-trypsin	1ezx	A	A	1.3E+01	2.9E+00	1.8E+01	4.0E+00
Trypsin	1ezx	B	B	1.3E+01	3.3E+00	1.8E+01	4.6E+00
Natural Killer Cell protease	1fi8	A	A	1.3E+01	1.8E+00	2.5E+01	3.4E+00
ecotin	1fi8	B	C D	1.3E+01	1.7E+00	2.5E+01	3.3E+00
Elastase	1fle	A	E	1.2E+01	1.5E+00	1.5E+01	1.9E+00
elafin	1fle	B	I	1.2E+01	1.3E+00	1.5E+01	1.6E+00
Caspase activator	1g73	A	A B	9.0E+00	9.2E-01	1.3E+01	1.3E+00
inhibitor	1g73	B	C	9.0E+00	1.1E+00	1.3E+01	1.6E+00
Kallikrein	1hia	A	A B	9.0E+00	1.2E+00	1.6E+01	2.1E+00
hirustasin	1hia	B	I	9.0E+00	9.4E-01	1.6E+01	1.7E+00
Factor Xa	1kig	A	H L	1.1E+01	1.3E+00	2.8E+01	3.3E+00
anti-coagulant	1kig	B	I	1.1E+01	1.2E+00	2.8E+01	3.1E+00
Trypsin	1mct	A	A	9.0E+00	1.4E+00	2.1E+01	3.2E+00
bitter gourd inhibitor	1mct	B	I	9.0E+00	1.1E+00	2.1E+01	2.5E+00
Mesenterocopeptidase	1mee	A	A	1.1E+01	1.4E+00	2.0E+01	2.5E+00
elgin C	1mee	B	I	1.1E+01	1.2E+00	2.0E+01	2.1E+00
Thrombin	1mkw	A	L H	1.3E+01	2.1E+00	1.7E+01	2.8E+00
prethrombin-2	1mkw	B	K	1.3E+01	2.0E+00	1.7E+01	2.6E+00

	pdbid			%non-pola	%polar	%charged
Protease-protease inhibitor						
Chymotrypsin	1acb	A	E	7.0E-01	2.9E-01	4.5E-03
elgin C	1acb	B	I	6.9E-01	2.1E-01	9.4E-02
Thrombin	1avg	A	H L	4.9E-01	3.0E-01	2.1E-01
triabin	1avg	B	I	6.7E-01	2.0E-01	1.3E-01
Trypsin	1avw	A	A	5.6E-01	4.3E-01	1.2E-02
soybean trypsin inhibitor	1avw	B	B	6.0E-01	1.7E-01	2.3E-01
Collagenase	1azz	A	A B	6.1E-01	3.7E-01	2.1E-02
ecotin	1azz	B	C	6.6E-01	2.6E-01	8.6E-02
Trypsin var. D189G, G226D	1brc	A	E	6.1E-01	3.7E-01	2.3E-02
amyloid beta-protein precursor	1brc	B	I	5.8E-01	2.6E-01	1.6E-01
Thrombin	1bth	A	L H	5.9E-01	3.4E-01	6.9E-02
bovine pancreatic trypsin inhibitor	1bth	B	P	5.5E-01	2.2E-01	2.3E-01
Trypsin	1c9t	A	A	6.0E-01	3.9E-01	3.9E-03
Bdellstasin	1c9t	B	G	7.0E-01	2.1E-01	9.7E-02
Chymotrypsin	1ca0	A	F G H	6.3E-01	3.7E-01	0.0E+00
amyloid beta-protein precursor	1ca0	B	I	4.9E-01	3.0E-01	2.2E-01
Chymotrypsin	1cbw	A	F G H	6.7E-01	3.3E-01	2.5E-04
Bovine pancreatic trypsin inhibitor	1cbw	B	I	6.0E-01	2.0E-01	1.9E-01
Chymotrypsin	1cgi	A	E	6.8E-01	3.0E-01	2.0E-02
trypsin inhibitor var 3	1cgi	B	I	6.7E-01	2.0E-01	1.2E-01
Chymotrypsin	1cgj	A	E	7.0E-01	2.8E-01	2.4E-02
trypsin inhibitor var 4	1cgj	B	I	6.9E-01	1.8E-01	1.3E-01
Chymotrypsin	1cho	A	E	6.8E-01	2.8E-01	3.5E-02
turkey ovomucoid	1cho	B	I	6.3E-01	2.4E-01	1.3E-01
Subtilisin Carlsberg	1cse	A	E	5.5E-01	4.5E-01	0.0E+00
elgin C	1cse	B	I	6.9E-01	1.6E-01	1.5E-01
Factor VIIa	1dan	A	L H	5.9E-01	1.9E-01	2.3E-01
soluble tissue factor	1dan	B	T U	5.0E-01	2.3E-01	2.7E-01
Elastase	1eai	A	B	5.8E-01	3.2E-01	9.1E-02
Chymotrypsin/elastase inhibitor	1eai	B	D	6.0E-01	3.1E-01	9.5E-02
ecotin	1ezs	A	A B	5.8E-01	2.0E-01	2.2E-01
Trypsin II anionic	1ezs	B	C	6.0E-01	2.7E-01	1.3E-01
anti-trypsin	1ezx	A	A	5.6E-01	2.9E-01	1.5E-01
Trypsin	1ezx	B	B	4.9E-01	4.7E-01	3.4E-02
Natural Killer Cell protease	1fi8	A	A	5.9E-01	3.7E-01	4.4E-02
ecotin	1fi8	B	C D	5.8E-01	2.5E-01	1.6E-01
Elastase	1fle	A	E	5.4E-01	2.2E-01	2.3E-01
elafin	1fle	B	I	7.0E-01	2.5E-01	5.9E-02
Caspase activator	1g73	A	A B	6.4E-01	2.3E-01	1.3E-01
inhibitor	1g73	B	C	6.0E-01	2.9E-01	1.1E-01
Kallikrein	1hia	A	A B	6.0E-01	3.8E-01	2.3E-02
hirustasin	1hia	B	I	6.0E-01	1.6E-01	2.4E-01
Factor Xa	1kig	A	H L	4.6E-01	2.9E-01	2.6E-01
anti-coagulant	1kig	B	I	5.5E-01	2.5E-01	2.0E-01
Trypsin	1mct	A	A	5.7E-01	4.2E-01	7.2E-03
bitter gourd inhibitor	1mct	B	I	6.3E-01	2.0E-01	1.7E-01
Mesenterocopeptidase	1mee	A	A	5.9E-01	3.8E-01	3.7E-02
elgin C	1mee	B	I	7.0E-01	1.9E-01	1.0E-01
Thrombin	1mkw	A	L H	7.0E-01	2.3E-01	6.1E-02
prethrombin-2	1mkw	B	K	6.4E-01	2.4E-01	1.3E-01

	pdbid			Reso-lution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Protease-protease inhibitor								
Trypsin	1ppe	A	E	2.8	23269	39500	9008	7.6E+02
cucurbita maxima	1ppe	B	I	2.8	3266	6585	2478	9.0E+02
Human leukocyte elastase	1ppf	A	E	2.4	23277	41190	10100	5.8E+02
turkey ovomucoid	1ppf	B	I	2.4	6008	11550	3909	7.3E+02
Papain	1stf	A	E	2.5	23396	39610	9289	8.0E+02
Stefin B	1stf	B	I	2.5	11135	20670	5974	9.6E+02
Trypsin	1tab	A	E	2.6	23269	39330	9074	5.7E+02
Bowman Bark inhibitor	1tab	B	I	2.6	4055	8295	3354	7.8E+02
Thrombin	1tbq	A	L H	3.2	35392	61310	14100	1.7E+03
rhodniin	1tbq	B	R	3.2	11046	21020	7301	1.8E+03
Thermitase	1tec	A	E	2.3	28323	45970	9483	6.5E+02
elgin C	1tec	B	I	2.3	7315	13630	4056	8.7E+02
Trypsin	1tfx	A	B	2.7	23440	39930	9256	5.8E+02
tissue factore pathway inhibitor	1tfx	B	D	2.7	6830	12700	4005	7.6E+02
Trypsinogen	1tgs	A	Z	2.2	23511	40260	9285	7.9E+02
PPTI	1tgs	B	I	2.2	6011	11760	4069	9.1E+02
Thrombin	1toc	A	A B	2.9	33014	56960	13220	1.7E+03
Ornithodorin	1toc	B	R	2.9	12686	23790	7987	1.8E+03
Anhydro trypsin	1tpa	A	E	2.1	23252	39540	9069	6.7E+02
PTI	1tpa	B	I	2.1	6510	12250	3886	7.7E+02
Kallikrein	2kai	A	A B	3	25534	43560	10430	6.4E+02
BPTI	2kai	B	I	3	6281	11810	3803	7.8E+02
beta-Trypsin	2ptc	A	E	2.9	23269	39560	9069	6.6E+02
PTI	2ptc	B	I	2.9	6510	12180	3865	7.6E+02
Subtilisin Carlsberg	2sic	A	E	2.65	27488	45500	9797	7.1E+02
streptomyces inhibitor	2sic	B	I	2.65	10912	20090	5857	8.8E+02
Subtilisin Carlsberg	2sni	A	E	2	27488	45580	9845	7.3E+02
chymotrypsin inhibitor	2sni	B	I	2	7269	13940	4314	8.8E+02
Proteinase B streptomyces B	3sgb	A	E	3	18631	31220	7366	5.9E+02
turkey ovomucoid	3sgb	B	I	3	5468	10360	3468	6.7E+02
Trypsinogen	3tpi	A	Z	3.2	23269	39930	9439	6.5E+02
PTI	3tpi	B	I	3.2	6510	12160	3874	7.6E+02
Thrombin	4htc	A	L H	2.7	32932	57110	13610	1.5E+03
Hirudin	4htc	B	I	2.7	6434	12820	4835	1.7E+03
Antibody-antigen								
HIV-1 capsid	1afv	A	A B	2.65	47320	80330	19060	7.0E+02
Fab 25.3	1afv	B	L H	2.65	33342	60040	8110	6.9E+02
Ab Fv	1ar1	A	C D	2.3	25006	42640	9964	6.5E+02
Cytochrome C	1ar1	B	A B	2.3	86930	141900	26990	6.6E+02
Fab	1bj1	A	H L	2.8	23290	39400	20090	9.1E+02
Vascular endothelial GF	1bj1	B	V W	2.8	21892	73830	13720	1.1E+03
Fab Hyhel-5	1bql	A	L H	2.5	46043	79750	19320	7.9E+02
Bobtail quail lysozyme	1bql	B	Y	2.5	14254	24730	6287	8.2E+02
Fab Hulys11	1bvk	A	D E	2.8	24720	42510	10310	6.4E+02
Lysozyme	1bvk	B	F	2.8	14297	24950	6543	6.5E+02
Fab Hyhel-63	1dqj	A	A B	2.5	46023	78860	18710	8.5E+02
Lysozyme	1dqj	B	B	2.5	14297	24640	6478	9.0E+02
Fv D1.3	1dvf	A	A B	2.3	24498	41550	9689	7.9E+02
Fv DE5.2	1dvf	B	C D	2.3	25028	43070	10230	8.3E+02

	pdbid			Circularity	RMSD to plane	Atom burial	Atom accessibility
Protease-protease inhibitor							
Trypsin	1ppe	A	E	5.3E-01	3.0E+00	6.3E+00	9.5E+00
cucurbita maxima	1ppe	B	I	3.3E-01	2.9E+00	1.1E+01	1.5E+01
Human leukocyte elastase	1ppf	A	E	7.2E-01	2.8E+00	7.0E+00	1.1E+01
turkey ovomucoid	1ppf	B	I	6.8E-01	2.3E+00	1.0E+01	1.6E+01
Papain	1stf	A	E	3.9E-01	3.6E+00	6.8E+00	1.1E+01
Stefin B	1stf	B	I	3.0E-01	2.0E+00	1.1E+01	1.7E+01
Trypsin	1tab	A	E	5.4E-01	2.7E+00	5.7E+00	8.4E+00
Bowman Bark inhibitor	1tab	B	I	2.1E-01	2.4E+00	1.4E+01	1.9E+01
Thrombin	1tbq	A	L H	2.2E-01	4.4E+00	7.9E+00	1.3E+01
rhodniin	1tbq	B	R	1.2E-01	4.2E+00	1.0E+01	1.5E+01
Thermitase	1tec	A	E	5.0E-01	2.6E+00	7.4E+00	1.0E+01
elgin C	1tec	B	I	3.5E-01	2.8E+00	1.2E+01	1.5E+01
Trypsin	1tfx	A	B	6.5E-01	2.8E+00	6.0E+00	9.8E+00
tissue factore pathway inhibitor	1tfx	B	D	3.7E-01	2.0E+00	1.2E+01	1.7E+01
Trypsinogen	1tgs	A	Z	5.1E-01	3.2E+00	7.0E+00	1.1E+01
PPTI	1tgs	B	I	2.5E-01	3.4E+00	1.1E+01	1.7E+01
Thrombin	1toc	A	A B	2.9E-01	4.4E+00	8.2E+00	1.3E+01
Ornithodorin	1toc	B	R	1.4E-01	4.1E+00	9.8E+00	1.5E+01
Anhydro trypsin	1tpa	A	E	4.6E-01	2.7E+00	6.2E+00	9.7E+00
PTI	1tpa	B	I	2.9E-01	1.9E+00	1.3E+01	1.6E+01
Kallikrein	2kai	A	A B	5.2E-01	3.4E+00	6.3E+00	9.7E+00
BPTI	2kai	B	I	2.9E-01	2.5E+00	1.3E+01	1.6E+01
beta-Trypsin	2ptc	A	E	4.8E-01	2.6E+00	6.5E+00	9.9E+00
PTI	2ptc	B	I	3.2E-01	1.9E+00	1.2E+01	1.7E+01
Subtilisin Carlsberg	2sic	A	E	4.7E-01	3.1E+00	6.6E+00	1.1E+01
streptomyces inhibitor	2sic	B	I	1.9E-01	2.4E+00	1.1E+01	1.5E+01
Subtilisin Carlsberg	2sni	A	E	5.3E-01	3.1E+00	7.0E+00	1.1E+01
chymotrypsin inhibitor	2sni	B	I	4.1E-01	3.3E+00	1.1E+01	1.6E+01
Proteinase B streptomyces B	3sgb	A	E	6.2E-01	2.6E+00	7.0E+00	1.1E+01
turkey ovomucoid	3sgb	B	I	3.2E-01	2.5E+00	9.7E+00	1.5E+01
Trypsinogen	3tpi	A	Z	4.5E-01	2.7E+00	6.4E+00	9.6E+00
PTI	3tpi	B	I	3.1E-01	1.9E+00	1.2E+01	1.6E+01
Thrombin	4htc	A	L H	2.4E-01	3.6E+00	7.7E+00	1.3E+01
Hirudin	4htc	B	I	1.2E-01	3.1E+00	1.1E+01	1.7E+01
Antibody-antigen							
HIV-1 capsid	1afv	A	A B	9.2E-01	2.6E+00	8.2E+00	1.3E+01
Fab 25.3	1afv	B	L H	3.9E-01	2.2E+00	9.2E+00	1.4E+01
Ab Fv	1ar1	A	C D	2.1E-01	1.6E+00	8.2E+00	1.4E+01
Cytochrome C	1ar1	B	A B	1.9E-01	1.1E+00	8.5E+00	1.4E+01
Fab	1bj1	A	H L	1.2E-01	4.7E+00	7.2E+00	1.0E+01
Vascular endothelial GF	1bj1	B	V W	3.0E-01	5.9E+00	1.0E+01	1.4E+01
Fab Hyhel-5	1bql	A	L H	6.0E-01	2.6E+00	8.4E+00	1.4E+01
Bobtail quail lysozyme	1bql	B	Y	5.4E-01	2.5E+00	9.9E+00	1.6E+01
Fab Hulys11	1bvk	A	D E	5.6E-01	1.7E+00	8.1E+00	1.3E+01
Lysozyme	1bvk	B	F	7.5E-01	1.6E+00	9.0E+00	1.5E+01
Fab Hyhel-63	1dqj	A	A B	3.6E-01	2.4E+00	8.1E+00	1.3E+01
Lysozyme	1dqj	B	B	3.8E-01	2.4E+00	9.3E+00	1.3E+01
Fv D1.3	1dvf	A	A B	7.5E-01	2.2E+00	7.9E+00	1.4E+01
Fv DE5.2	1dvf	B	C D	6.5E-01	2.1E+00	1.0E+01	1.6E+01

	pdbid			Shape comple- mentarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Protease-protease inhibitor						
Trypsin	1ppe	A	E	7.5E-01	1.0E+00	1.5E+00
cucurbita maxima	1ppe	B	I	7.5E-01	1.2E+00	2.0E+00
Human leukocyte elastase	1ppf	A	E	7.7E-01	1.0E+00	1.9E+00
turkey ovomucoid	1ppf	B	I	7.7E-01	1.1E+00	1.7E+00
Papain	1stf	A	E	6.6E-01	1.1E+00	1.7E+00
Stefin B	1stf	B	I	6.6E-01	1.1E+00	1.9E+00
Trypsin	1tab	A	E	6.8E-01	9.9E-01	1.6E+00
Bowman Bark inhibitor	1tab	B	I	6.8E-01	1.3E+00	2.6E+00
Thrombin	1tbq	A	L H	6.9E-01	1.0E+00	2.0E+00
rhodniin	1tbq	B	R	6.9E-01	1.3E+00	2.0E+00
Thermitase	1tec	A	E	7.5E-01	1.2E+00	1.7E+00
elgin C	1tec	B	I	7.5E-01	1.2E+00	1.8E+00
Trypsin	1tfx	A	B	7.4E-01	1.0E+00	1.6E+00
tissue factore pathway inhibitor	1tfx	B	D	7.4E-01	1.1E+00	2.1E+00
Trypsinogen	1tgs	A	Z	7.5E-01	1.0E+00	1.7E+00
PPTI	1tgs	B	I	7.5E-01	1.0E+00	2.0E+00
Thrombin	1toc	A	A B	6.5E-01	1.0E+00	2.0E+00
Ornithodorin	1toc	B	R	6.5E-01	1.1E+00	2.2E+00
Anhydro trypsin	1tpa	A	E	7.6E-01	1.0E+00	1.6E+00
PTI	1tpa	B	I	7.6E-01	1.2E+00	1.9E+00
Kallikrein	2kai	A	A B	7.4E-01	1.1E+00	1.9E+00
BPTI	2kai	B	I	7.4E-01	1.3E+00	2.0E+00
beta-Trypsin	2ptc	A	E	7.7E-01	9.8E-01	1.5E+00
PTI	2ptc	B	I	7.7E-01	1.2E+00	1.9E+00
Subtilisin Carlsberg	2sic	A	E	7.4E-01	1.2E+00	1.6E+00
streptomyces inhibitor	2sic	B	I	7.4E-01	1.0E+00	1.7E+00
Subtilisin Carlsberg	2sni	A	E	7.2E-01	1.3E+00	1.8E+00
chymotrypsin inhibitor	2sni	B	I	7.2E-01	1.1E+00	2.1E+00
Proteinase B streptomyces B	3sgb	A	E	7.9E-01	1.1E+00	1.5E+00
turkey ovomucoid	3sgb	B	I	7.9E-01	9.6E-01	1.7E+00
Trypsinogen	3tpi	A	Z	7.7E-01	1.1E+00	1.5E+00
PTI	3tpi	B	I	7.7E-01	1.1E+00	1.8E+00
Thrombin	4htc	A	L H	6.5E-01	1.1E+00	1.7E+00
Hirudin	4htc	B	I	6.5E-01	1.4E+00	1.8E+00
Antibody-antigen						
HIV-1 capsid	1afv	A	A B	6.3E-01	1.0E+00	2.1E+00
Fab 25.3	1afv	B	L H	6.3E-01	1.0E+00	2.2E+00
Ab Fv	1ar1	A	C D	6.9E-01	1.0E+00	2.0E+00
Cytochrome C	1ar1	B	A B	6.9E-01	1.1E+00	2.2E+00
Fab	1bj1	A	H L	7.9E-01	1.1E+00	1.9E+00
Vascular endothelial GF	1bj1	B	V W	7.9E-01	1.0E+00	1.8E+00
Fab Hyhel-5	1bql	A	L H	7.0E-01	9.8E-01	1.8E+00
Bobtail quail lysozyme	1bql	B	Y	7.0E-01	1.1E+00	1.7E+00
Fab Hulys11	1bvk	A	D E	6.9E-01	9.9E-01	1.8E+00
Lysozyme	1bvk	B	F	6.9E-01	9.1E-01	1.8E+00
Fab Hyhel-63	1dqj	A	A B	7.0E-01	1.1E+00	1.6E+00
Lysozyme	1dqj	B	B	7.0E-01	1.1E+00	1.5E+00
Fv D1.3	1dvf	A	A B	7.5E-01	9.5E-01	1.7E+00
Fv DE5.2	1dvf	B	C D	7.5E-01	9.5E-01	1.7E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
Protease-protease inhibitor							
Trypsin	1ppe	A	E	2.0E+00	2.6E-01	9.0E+00	1.2E+00
cucurbita maxima	1ppe	B	I	2.0E+00	2.2E-01	9.0E+00	1.0E+00
Human leukocyte elastase	1ppf	A	E	1.9E+01	3.3E+00	4.2E+01	7.3E+00
turkey ovomucoid	1ppf	B	I	1.9E+01	2.6E+00	4.2E+01	5.7E+00
Papain	1stf	A	E	1.2E+01	1.5E+00	2.0E+01	2.5E+00
Stefin B	1stf	B	I	1.2E+01	1.3E+00	2.0E+01	2.1E+00
Trypsin	1tab	A	E	1.2E+01	2.1E+00	1.4E+01	2.4E+00
Bowman Bark inhibitor	1tab	B	I	1.2E+01	1.5E+00	1.4E+01	1.8E+00
Thrombin	1tbq	A	L H	1.5E+01	9.1E-01	2.0E+01	1.2E+00
rhodniin	1tbq	B	R	1.5E+01	8.3E-01	2.0E+01	1.1E+00
Thermitase	1tec	A	E	5.0E+00	7.7E-01	9.0E+00	1.4E+00
elgin C	1tec	B	I	5.0E+00	5.8E-01	9.0E+00	1.0E+00
Trypsin	1tfx	A	B	7.0E+00	1.2E+00	2.0E+01	3.4E+00
tissue factore pathway inhibitor	1tfx	B	D	7.0E+00	9.2E-01	2.0E+01	2.6E+00
Trypsinogen	1tgs	A	Z	1.6E+01	2.0E+00	2.7E+01	3.4E+00
PPTI	1tgs	B	I	1.6E+01	1.8E+00	2.7E+01	3.0E+00
Thrombin	1toc	A	A B	7.0E+00	4.1E-01	1.7E+01	1.0E+00
Ornithodorin	1toc	B	R	7.0E+00	4.0E-01	1.7E+01	9.6E-01
Anhydro trypsin	1tpa	A	E	3.4E+01	5.1E+00	5.2E+01	7.8E+00
PTI	1tpa	B	I	3.4E+01	4.4E+00	5.2E+01	6.8E+00
Kallikrein	2kai	A	A B	1.8E+01	2.8E+00	3.1E+01	4.8E+00
BPTI	2kai	B	I	1.8E+01	2.3E+00	3.1E+01	4.0E+00
beta-Trypsin	2ptc	A	E	2.5E+01	3.8E+00	5.5E+01	8.3E+00
PTI	2ptc	B	I	2.5E+01	3.3E+00	5.5E+01	7.2E+00
Subtilisin Carlsberg	2sic	A	E	1.3E+01	1.8E+00	2.8E+01	3.9E+00
streptomyces inhibitor	2sic	B	I	1.3E+01	1.5E+00	2.8E+01	3.2E+00
Subtilisin Carlsberg	2sni	A	E	1.1E+01	1.5E+00	2.1E+01	2.9E+00
chymotrypsin inhibitor	2sni	B	I	1.1E+01	1.3E+00	2.1E+01	2.4E+00
Proteinase B streptomyces B	3sgb	A	E	1.8E+01	3.1E+00	3.0E+01	5.1E+00
turkey ovomucoid	3sgb	B	I	1.8E+01	2.7E+00	3.0E+01	4.5E+00
Trypsinogen	3tpi	A	Z	6.0E+00	9.2E-01	1.6E+01	2.5E+00
PTI	3tpi	B	I	6.0E+00	7.8E-01	1.6E+01	2.1E+00
Thrombin	4htc	A	L H	9.0E+00	5.8E-01	1.3E+01	8.4E-01
Hirudin	4htc	B	I	9.0E+00	5.2E-01	1.3E+01	7.5E-01
Antibody-antigen							
HIV-1 capsid	1afv	A	A B	2.1E+01	3.0E+00	3.7E+01	5.3E+00
Fab 25.3	1afv	B	L H	2.1E+01	3.1E+00	3.7E+01	5.4E+00
Ab Fv	1ar1	A	C D	1.8E+01	2.8E+00	3.1E+01	4.7E+00
Cytochrome C	1ar1	B	A B	1.8E+01	2.7E+00	3.1E+01	4.7E+00
Fab	1bj1	A	H L	1.4E+01	1.5E+00	2.7E+01	3.0E+00
Vascular endothelial GF	1bj1	B	V W	1.4E+01	1.3E+00	2.7E+01	2.5E+00
Fab Hyhel-5	1bql	A	L H	2.0E+01	2.5E+00	3.7E+01	4.7E+00
Bobtail quail lysozyme	1bql	B	Y	2.0E+01	2.5E+00	3.7E+01	4.5E+00
Fab Hulys11	1bvk	A	D E	6.0E+00	9.4E-01	1.3E+01	2.0E+00
Lysozyme	1bvk	B	F	6.0E+00	9.2E-01	1.3E+01	2.0E+00
Fab Hyhel-63	1dqj	A	A B	1.1E+01	1.3E+00	1.8E+01	2.1E+00
Lysozyme	1dqj	B	B	1.1E+01	1.2E+00	1.8E+01	2.0E+00
Fv D1.3	1dvf	A	A B	2.0E+00	2.5E-01	7.0E+00	8.9E-01
Fv DE5.2	1dvf	B	C D	2.0E+00	2.4E-01	7.0E+00	8.4E-01

	pdbid			%non-pola	%polar	%charged
Protease-protease inhibitor						
Trypsin	1ppe	A	E	5.5E-01	4.2E-01	3.0E-02
cucurbita maxima	1ppe	B	I	6.2E-01	1.8E-01	2.0E-01
Human leukocyte elastase	1ppf	A	E	7.4E-01	2.7E-01	0.0E+00
turkey ovomucoid	1ppf	B	I	7.3E-01	2.2E-01	5.5E-02
Papain	1stf	A	E	5.9E-01	4.0E-01	7.6E-03
Stefin B	1stf	B	I	7.0E-01	2.9E-01	1.8E-02
Trypsin	1tab	A	E	5.3E-01	4.4E-01	3.8E-02
Bowman Bark inhibitor	1tab	B	I	5.2E-01	2.9E-01	1.9E-01
Thrombin	1tbq	A	L H	5.8E-01	2.0E-01	2.2E-01
rhodniin	1tbq	B	R	6.0E-01	2.6E-01	1.4E-01
Thermitase	1tec	A	E	5.6E-01	4.4E-01	0.0E+00
elgin C	1tec	B	I	7.0E-01	1.8E-01	1.2E-01
Trypsin	1tfx	A	B	5.7E-01	4.1E-01	2.3E-02
tissue factore pathway inhibitor	1tfx	B	D	5.9E-01	2.4E-01	1.7E-01
Trypsinogen	1tgs	A	Z	5.7E-01	4.2E-01	6.6E-03
PPTI	1tgs	B	I	7.0E-01	2.4E-01	6.7E-02
Thrombin	1toc	A	A B	4.9E-01	2.9E-01	2.3E-01
Ornithodorin	1toc	B	R	6.4E-01	2.7E-01	9.5E-02
Anhydro trypsin	1tpa	A	E	5.8E-01	4.2E-01	4.5E-03
PTI	1tpa	B	I	5.6E-01	2.0E-01	2.4E-01
Kallikrein	2kai	A	A B	6.5E-01	3.3E-01	2.3E-02
BPTI	2kai	B	I	5.7E-01	2.3E-01	2.0E-01
beta-Trypsin	2ptc	A	E	5.7E-01	4.2E-01	6.5E-03
PTI	2ptc	B	I	5.5E-01	2.0E-01	2.5E-01
Subtilisin Carlsberg	2sic	A	E	6.3E-01	3.3E-01	4.4E-02
streptomyces inhibitor	2sic	B	I	6.1E-01	3.1E-01	8.3E-02
Subtilisin Carlsberg	2sni	A	E	5.6E-01	3.9E-01	5.7E-02
chymotrypsin inhibitor	2sni	B	I	7.4E-01	1.7E-01	9.1E-02
Proteinase B streptomyces B	3sgb	A	E	6.2E-01	3.2E-01	5.4E-02
turkey ovomucoid	3sgb	B	I	6.9E-01	2.0E-01	1.0E-01
Trypsinogen	3tpi	A	Z	5.9E-01	4.0E-01	7.1E-03
PTI	3tpi	B	I	5.4E-01	2.0E-01	2.6E-01
Thrombin	4htc	A	L H	5.3E-01	2.2E-01	2.5E-01
Hirudin	4htc	B	I	6.1E-01	2.7E-01	1.2E-01
Antibody-antigen						
HIV-1 capsid	1afv	A	A B	5.9E-01	3.5E-01	6.0E-02
Fab 25.3	1afv	B	L H	5.5E-01	1.0E-01	3.5E-01
Ab Fv	1ar1	A	C D	4.5E-01	4.1E-01	1.4E-01
Cytochrome C	1ar1	B	A B	5.7E-01	2.2E-01	2.2E-01
Fab	1bj1	A	H L	6.0E-01	4.0E-01	0.0E+00
Vascular endothelial GF	1bj1	B	V W	5.1E-01	4.2E-01	6.4E-02
Fab Hyhel-5	1bql	A	L H	5.4E-01	3.2E-01	1.4E-01
Bobtail quail lysozyme	1bql	B	Y	5.5E-01	3.2E-01	1.4E-01
Fab Hulys11	1bvk	A	D E	5.0E-01	3.2E-01	1.8E-01
Lysozyme	1bvk	B	F	5.1E-01	3.6E-01	1.3E-01
Fab Hyhel-63	1dqj	A	A B	4.6E-01	4.9E-01	4.4E-02
Lysozyme	1dqj	B	B	5.0E-01	2.8E-01	2.3E-01
Fv D1.3	1dvf	A	A B	4.7E-01	3.4E-01	1.9E-01
Fv DE5.2	1dvf	B	C D	4.5E-01	4.1E-01	1.4E-01

Antibody-antigen	pdbid			Reso-lution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Scfv 1f9	1dzb	A	A B	2.1	24456	42550	10350	7.7E+02
turkey lysozyme	1dzb	B	X	2.1	14193	24590	6358	8.9E+02
Fab	1e6j	A	H L	2.5	46457	81490	19730	5.7E+02
Capsid protein p24	1e6j	B	P	2.5	23339	44140	13290	6.7E+02
Fab	1eo8	A	H L	2.5	46385	78850	18920	7.6E+02
Hemaglutinin	1eo8	B	A B	2.5	55146	94560	23160	7.6E+02
Fab F9.13.7	1fbi	A	L H	2.5	46289	80930	19680	8.5E+02
Lysozyme	1fbi	B	X	2.5	14292	25070	6667	8.3E+02
Fab nmc-4	1fns	A	L H	2.6	47188	80950	19400	6.1E+02
von Willebrand factor	1fns	B	A	2.6	22361	39150	9311	6.7E+02
Kappa Fab	1fsk	A	K L	2	47445	81120	19300	7.3E+02
major pollen antigen bet v 1-a	1fsk	B	J	2	17364	31270	8323	8.7E+02
Fab 17b	1gc1	A	H L	2.5	47694	81220	18780	5.6E+02
gp120	1gc1	B	G C	2.5	52858	92170	9413	6.0E+02
Fab 730.1.4	1iai	A	L H	2.8	47386	83050	19590	9.3E+02
Fab 409.5.3	1iai	B	M I	2.8	47252	81990	19190	9.4E+02
Fab	1jhl	A	L H	2.3	24780	42470	10120	5.9E+02
protein G domain III	1jhl	B	A	2.3	14238	24730	6364	6.4E+02
Fab A6	1jrh	A	L H	2.6	37906	67800	17710	7.5E+02
interferon gamma receptor alpha cha	1jrh	B	I	2.6	10770	20360	6045	7.9E+02
Fab desire-1	1kb5	A	L H	2.2	47186	80360	18840	1.1E+03
Kb5-c20 T-cell receptor	1kb5	B	A B	2.2	26110	45150	11090	1.2E+03
Lysozyme	1mel	A	L	2	14027	23920	6137	7.9E+02
VH single domain antibody	1mel	B	A	2	13854	24340	6413	8.9E+02
Lysozyme	1mlc	A	E	2.3	14297	24460	6358	7.2E+02
Fab D44.1	1mlc	B	A B	2.3	46704	80120	19090	6.5E+02
Fab	1nca	A	L H	2.3	47222	81220	19170	9.2E+02
N9 neuraminidase-nc41	1nca	B	N	2.3	43720	71700	14820	9.8E+02
N15 α-β T-cell receptor	1nfd	A	A B	1.9	47226	82100	20900	7.5E+02
Fab h57	1nfd	B	E F	1.9	49772	86580	21240	8.6E+02
Fab NC-10	1nmb	A	L H	2	25491	43730	10740	6.4E+02
N9-neuraminidase	1nmb	B	N	2	43528	71060	14570	6.9E+02
Kappa Fab	1nsn	A	L H	2.2	46550	80060	19110	8.8E+02
staphylococcal nuclease	1nsn	B	S	2.2	15703	28740	7966	8.8E+02
Fab 184.1	1osp	A	L H	2.6	47024	80940	19510	7.6E+02
protein A outer surface	1osp	B	O	2.6	26940	50150	14150	6.9E+02
Fab	1qfw	A	HLIM	2.3	48647	86710	22180	1.5E+03
Gonadotropin	1qfw	B	A B	2.3	21310	39700	11420	1.7E+03
rhinovirus coat protein	1rvf	A	VWXY	3.5	88839	151400	31330	9.1E+02
Fab 17-1a	1rvf	B	H L	3.5	24831	40570	9697	8.8E+02
Fv D1.3	1vfb	A	A B	2.8	24498	42110	9966	6.5E+02
Lysozyme	1vfb	B	C	2.8	14297	24880	6438	7.1E+02
Fab e8	1wej	A	H L	2.7	47380	81080	19260	5.3E+02
cytochrome C	1wej	B	F	2.7	11676	21560	6139	6.1E+02
lambda Fab	2vir	A	A B	3.95	46259	79760	19280	6.5E+02
hemagglutinin	2vir	B	C	3.95	29449	49800	11610	6.0E+02
hyhel-5	3hfl	A	L H	2	46043	79430	19030	8.4E+02
Lysozyme	3hfl	B	Y	2	14297	24660	6286	8.5E+02
hyhel-10	3hfm	A	L H	3.2	46758	81600	19440	7.4E+02
Lysozyme	3hfm	B	Y	3.2	14297	24910	6560	8.3E+02

Antibody-antigen	pdbid			Circularity	RMSD to plane	Atom burial	Atom access-ibility
Scfv 1f9	1dzb	A	A B	5.2E-01	2.5E+00	7.0E+00	1.1E+01
turkey lysozyme	1dzb	B	X	5.1E-01	2.6E+00	1.0E+01	1.5E+01
Fab	1e6j	A	H L	6.9E-01	2.1E+00	6.1E+00	1.0E+01
Capsid protein p24	1e6j	B	P	7.8E-01	2.1E+00	1.1E+01	1.7E+01
Fab	1eo8	A	H L	4.9E-01	2.1E+00	8.9E+00	1.4E+01
Hemaglutinin	1eo8	B	A B	4.4E-01	2.2E+00	8.4E+00	1.3E+01
Fab F9.13.7	1fbi	A	L H	5.2E-01	2.1E+00	9.1E+00	1.3E+01
Lysozyme	1fbi	B	X	4.9E-01	1.9E+00	8.1E+00	1.3E+01
Fab nmc-4	1fns	A	L H	5.1E-01	2.3E+00	7.5E+00	1.1E+01
von Willebrand factor	1fns	B	A	7.1E-01	2.3E+00	1.1E+01	1.7E+01
Kappa Fab	1fsk	A	K L	8.4E-01	2.6E+00	6.2E+00	1.1E+01
major pollen antigen bet v 1-a	1fsk	B	J	8.2E-01	2.6E+00	1.0E+01	1.4E+01
Fab 17b	1gc1	A	H L	3.3E-01	1.7E+00	8.2E+00	1.3E+01
gp120	1gc1	B	G C	3.8E-01	1.7E+00	1.0E+01	1.3E+01
Fab 730.1.4	1iai	A	L H	3.5E-01	2.7E+00	8.8E+00	1.3E+01
Fab 409.5.3	1iai	B	M I	3.3E-01	2.9E+00	8.5E+00	1.4E+01
Fab	1jhl	A	L H	6.2E-01	1.7E+00	7.2E+00	1.3E+01
protein G domain III	1jhl	B	A	5.3E-01	1.8E+00	8.8E+00	1.5E+01
Fab A6	1jrh	A	L H	4.4E-01	2.5E+00	8.1E+00	1.3E+01
interferon gamma receptor alpha cha	1jrh	B	I	6.3E-01	2.3E+00	1.0E+01	1.5E+01
Fab desire-1	1kb5	A	L H	4.8E-01	3.0E+00	9.3E+00	1.5E+01
Kb5-c20 T-cell receptor	1kb5	B	A B	7.7E-01	2.9E+00	9.2E+00	1.4E+01
Lysozyme	1mel	A	L	6.1E-01	3.3E+00	8.3E+00	1.3E+01
VH single domain antibody	1mel	B	A	9.0E-01	3.4E+00	1.0E+01	1.4E+01
Lysozyme	1mlc	A	E	5.2E-01	2.0E+00	1.0E+01	1.5E+01
Fab D44.1	1mlc	B	A B	6.2E-01	1.7E+00	7.6E+00	1.1E+01
Fab	1nca	A	L H	6.5E-01	2.4E+00	7.5E+00	1.3E+01
N9 neuraminidase-nc41	1nca	B	N	4.7E-01	2.2E+00	9.7E+00	1.6E+01
N15 α-β T-cell receptor	1nfd	A	A B	4.6E-01	3.0E+00	7.2E+00	1.4E+01
Fab h57	1nfd	B	E F	7.6E-01	3.3E+00	1.1E+01	1.5E+01
Fab NC-10	1nmb	A	L H	2.3E-01	1.6E+00	8.0E+00	1.3E+01
N9-neuraminidase	1nmb	B	N	2.3E-01	1.5E+00	1.0E+01	1.6E+01
Kappa Fab	1nsn	A	L H	4.8E-01	3.4E+00	9.0E+00	1.7E+01
staphylococcal nuclease	1nsn	B	S	7.5E-01	3.5E+00	8.5E+00	1.5E+01
Fab 184.1	1osp	A	L H	2.6E-01	1.7E+00	9.2E+00	1.4E+01
protein A outer surface	1osp	B	O	2.7E-01	1.6E+00	8.5E+00	1.5E+01
Fab	1qfw	A	HLIM	2.3E-02	4.0E+00	7.6E+00	1.3E+01
Gonadotropin	1qfw	B	A B	2.9E-02	3.3E+00	1.0E+01	1.6E+01
rhinovirus coat protein	1rvf	A	VWXY	1.8E-01	3.4E+00	9.4E+00	1.6E+01
Fab 17-1a	1rvf	B	H L	2.4E-01	3.5E+00	8.1E+00	1.3E+01
Fv D1.3	1vfb	A	A B	6.1E-01	1.6E+00	8.0E+00	1.3E+01
Lysozyme	1vfb	B	C	6.8E-01	1.7E+00	9.0E+00	1.4E+01
Fab e8	1wej	A	H L	2.6E-01	1.5E+00	7.2E+00	1.3E+01
cytochrome C	1wej	B	F	3.8E-01	1.6E+00	1.0E+01	1.6E+01
lambda Fab	2vir	A	A B	5.6E-01	1.8E+00	8.9E+00	1.4E+01
hemagglutinin	2vir	B	C	4.8E-01	2.4E+00	7.5E+00	1.3E+01
hyhel-5	3hfl	A	L H	6.3E-01	2.8E+00	7.5E+00	1.2E+01
Lysozyme	3hfl	B	Y	4.7E-01	3.0E+00	8.3E+00	1.3E+01
hyhel-10	3hfm	A	L H	3.8E-01	2.5E+00	8.0E+00	1.2E+01
Lysozyme	3hfm	B	Y	3.8E-01	2.4E+00	8.7E+00	1.3E+01

Antibody-antigen	pdbid			Shape comple- mentarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Scfv 1f9	1dzb	A	A B	6.8E-01	1.1E+00	1.7E+00
turkey lysozyme	1dzb	B	X	6.8E-01	1.0E+00	1.7E+00
Fab	1e6j	A	H L	7.4E-01	1.1E+00	1.9E+00
Capsid protein p24	1e6j	B	P	7.4E-01	1.3E+00	2.3E+00
Fab	1eo8	A	H L	7.3E-01	1.1E+00	2.0E+00
Hemaglutinin	1eo8	B	A B	7.3E-01	1.1E+00	1.8E+00
Fab F9.13.7	1fbi	A	L H	6.5E-01	1.2E+00	1.6E+00
Lysozyme	1fbi	B	X	6.5E-01	1.1E+00	1.9E+00
Fab nmc-4	1fns	A	L H	8.3E-01	1.2E+00	2.2E+00
von Willebrand factor	1fns	B	A	8.3E-01	1.5E+00	2.4E+00
Kappa Fab	1fsk	A	K L	7.0E-01	9.2E-01	2.0E+00
major pollen antigen bet v 1-a	1fsk	B	J	7.0E-01	1.0E+00	1.9E+00
Fab 17b	1gc1	A	H L	7.0E-01	1.1E+00	2.2E+00
gp120	1gc1	B	G C	7.0E-01	1.2E+00	2.6E+00
Fab 730.1.4	1iai	A	L H	6.2E-01	9.2E-01	2.0E+00
Fab 409.5.3	1iai	B	M I	6.2E-01	9.3E-01	2.0E+00
Fab	1jhl	A	L H	6.5E-01	1.0E+00	1.8E+00
protein G domain III	1jhl	B	A	6.5E-01	9.5E-01	1.9E+00
Fab A6	1jrh	A	L H	7.8E-01	1.1E+00	1.9E+00
interferon gamma receptor alpha cha	1jrh	B	I	7.8E-01	1.0E+00	2.1E+00
Fab desire-1	1kb5	A	L H	5.4E-01	1.0E+00	1.8E+00
Kb5-c20 T-cell receptor	1kb5	B	A B	5.4E-01	1.1E+00	1.7E+00
Lysozyme	1mel	A	L	7.8E-01	1.2E+00	1.9E+00
VH single domain antibody	1mel	B	A	7.8E-01	1.1E+00	1.8E+00
Lysozyme	1mlc	A	E	6.6E-01	1.2E+00	1.7E+00
Fab D44.1	1mlc	B	A B	6.6E-01	9.7E-01	1.8E+00
Fab	1nca	A	L H	6.5E-01	1.1E+00	2.1E+00
N9 neuraminidase-nc41	1nca	B	N	6.5E-01	1.1E+00	2.4E+00
N15 α-β T-cell receptor	1nfd	A	A B	6.7E-01	1.1E+00	2.0E+00
Fab h57	1nfd	B	E F	6.7E-01	8.7E-01	2.0E+00
Fab NC-10	1nmb	A	L H	6.6E-01	9.7E-01	1.6E+00
N9-neuraminidase	1nmb	B	N	6.6E-01	1.0E+00	1.8E+00
Kappa Fab	1nsn	A	L H	5.9E-01	1.1E+00	2.1E+00
staphylococcal nuclease	1nsn	B	S	5.9E-01	1.0E+00	2.2E+00
Fab 184.1	1osp	A	L H	7.7E-01	9.7E-01	1.5E+00
protein A outer surface	1osp	B	O	7.7E-01	1.1E+00	1.8E+00
Fab	1qfw	A	HLIM	5.9E-01	1.0E+00	1.9E+00
Gonadotropin	1qfw	B	A B	5.9E-01	1.1E+00	2.1E+00
rhinovirus coat protein	1rvf	A	VWXY	6.3E-01	1.2E+00	2.6E+00
Fab 17-1a	1rvf	B	H L	6.3E-01	1.2E+00	2.2E+00
Fv D1.3	1vfb	A	A B	7.3E-01	8.8E-01	1.4E+00
Lysozyme	1vfb	B	C	7.3E-01	9.2E-01	1.4E+00
Fab e8	1wej	A	H L	7.8E-01	9.8E-01	1.5E+00
cytochrome C	1wej	B	F	7.8E-01	1.3E+00	1.7E+00
lambda Fab	2vir	A	A B	6.6E-01	-9.0E-02	2.0E+00
hemagglutinin	2vir	B	C	6.6E-01	-3.0E-01	2.0E+00
hyhel-5	3hfl	A	L H	7.2E-01	1.0E+00	1.7E+00
Lysozyme	3hfl	B	Y	7.2E-01	1.0E+00	1.6E+00
hyhel-10	3hfm	A	L H	6.8E-01	1.0E+00	1.9E+00
Lysozyme	3hfm	B	Y	6.8E-01	1.1E+00	1.9E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstrom	Hbonds <4.0 A	per 100 sq angstrom
Antibody-antigen							
Scfv 1f9	1dzb	A	A B	9.0E+00	1.2E+00	1.6E+01	2.1E+00
turkey lysozyme	1dzb	B	X	9.0E+00	1.0E+00	1.6E+01	1.8E+00
Fab	1e6j	A	H L	9.0E+00	1.6E+00	1.6E+01	2.8E+00
Capsid protein p24	1e6j	B	P	9.0E+00	1.3E+00	1.6E+01	2.4E+00
Fab	1eo8	A	H L	7.0E+00	9.2E-01	1.9E+01	2.5E+00
Hemaglutinin	1eo8	B	A B	7.0E+00	9.3E-01	1.9E+01	2.5E+00
Fab F9.13.7	1fbi	A	L H	1.3E+01	1.5E+00	2.1E+01	2.5E+00
Lysozyme	1fbi	B	X	1.3E+01	1.6E+00	2.1E+01	2.5E+00
Fab nmc-4	1fns	A	L H	1.1E+01	1.8E+00	2.0E+01	3.3E+00
von Willebrand factor	1fns	B	A	1.1E+01	1.6E+00	2.0E+01	3.0E+00
Kappa Fab	1fsk	A	K L	9.0E+00	1.2E+00	2.0E+01	2.7E+00
major pollen antigen bet v 1-a	1fsk	B	J	9.0E+00	1.0E+00	2.0E+01	2.3E+00
Fab 17b	1gc1	A	H L	8.0E+00	1.4E+00	2.4E+01	4.3E+00
gp120	1gc1	B	G C	8.0E+00	1.3E+00	2.4E+01	4.0E+00
Fab 730.1.4	1iai	A	L H	1.5E+01	1.6E+00	2.6E+01	2.8E+00
Fab 409.5.3	1iai	B	M I	1.5E+01	1.6E+00	2.6E+01	2.8E+00
Fab	1jhl	A	L H	1.9E+01	3.2E+00	3.1E+01	5.3E+00
protein G domain III	1jhl	B	A	1.9E+01	3.0E+00	3.1E+01	4.8E+00
Fab A6	1jrh	A	L H	5.0E+00	6.7E-01	1.0E+01	1.3E+00
interferon gamma receptor alpha cha	1jrh	B	I	5.0E+00	6.3E-01	1.0E+01	1.3E+00
Fab desire-1	1kb5	A	L H	9.0E+00	8.0E-01	2.5E+01	2.2E+00
Kb5-c20 T-cell receptor	1kb5	B	A B	9.0E+00	7.8E-01	2.5E+01	2.2E+00
Lysozyme	1mel	A	L	1.2E+01	1.5E+00	2.6E+01	3.3E+00
VH single domain antibody	1mel	B	A	1.2E+01	1.4E+00	2.6E+01	2.9E+00
Lysozyme	1mlc	A	E	1.0E+01	1.4E+00	2.3E+01	3.2E+00
Fab D44.1	1mlc	B	A B	1.0E+01	1.5E+00	2.3E+01	3.5E+00
Fab	1nca	A	L H	1.0E+01	1.1E+00	1.9E+01	2.1E+00
N9 neuraminidase-nc41	1nca	B	N	1.0E+01	1.0E+00	1.9E+01	1.9E+00
N15 α-β T-cell receptor	1nfd	A	A B	1.4E+01	1.9E+00	2.7E+01	3.6E+00
Fab h57	1nfd	B	E F	1.4E+01	1.6E+00	2.7E+01	3.1E+00
Fab NC-10	1nmb	A	L H	1.0E+01	1.6E+00	2.5E+01	3.9E+00
N9-neuraminidase	1nmb	B	N	1.0E+01	1.5E+00	2.5E+01	3.6E+00
Kappa Fab	1nsn	A	L H	9.0E+00	1.0E+00	1.7E+01	1.9E+00
staphylococcal nuclease	1nsn	B	S	9.0E+00	1.0E+00	1.7E+01	1.9E+00
Fab 184.1	1osp	A	L H	3.0E+00	3.9E-01	5.0E+00	6.6E-01
protein A outer surface	1osp	B	O	3.0E+00	4.3E-01	5.0E+00	7.2E-01
Fab	1qfw	A	HLIM	2.0E+01	1.3E+00	4.0E+01	2.7E+00
Gonadotropin	1qfw	B	A B	2.0E+01	1.2E+00	4.0E+01	2.4E+00
rhinovirus coat protein	1rvf	A	VWXY	8.0E+00	8.8E-01	1.6E+01	1.8E+00
Fab 17-1a	1rvf	B	H L	8.0E+00	9.0E-01	1.6E+01	1.8E+00
Fv D1.3	1vfb	A	A B	1.2E+01	1.9E+00	1.6E+01	2.5E+00
Lysozyme	1vfb	B	C	1.2E+01	1.7E+00	1.6E+01	2.3E+00
Fab e8	1wej	A	H L	1.5E+01	2.8E+00	2.9E+01	5.5E+00
cytochrome C	1wej	B	F	1.5E+01	2.4E+00	2.9E+01	4.7E+00
lambda Fab	2vir	A	A B	1.0E+00	1.5E-01	2.0E+00	3.1E-01
hemagglutinin	2vir	B	C	1.0E+00	1.7E-01	2.0E+00	3.3E-01
hyhel-5	3hfl	A	L H	1.5E+01	1.8E+00	2.2E+01	2.6E+00
Lysozyme	3hfl	B	Y	1.5E+01	1.8E+00	2.2E+01	2.6E+00
hyhel-10	3hfm	A	L H	1.2E+01	1.6E+00	2.5E+01	3.4E+00
Lysozyme	3hfm	B	Y	1.2E+01	1.4E+00	2.5E+01	3.0E+00

	pdbid			%non-pola	%polar	%charged
Antibody-antigen						
Scfv 1f9	1dzb	A	A B	5.4E-01	3.3E-01	1.3E-01
turkey lysozyme	1dzb	B	X	5.2E-01	2.6E-01	2.2E-01
Fab	1e6j	A	H L	6.4E-01	3.6E-01	0.0E+00
Capsid protein p24	1e6j	B	P	7.1E-01	1.7E-01	1.3E-01
Fab	1eo8	A	H L	5.4E-01	3.6E-01	1.1E-01
Hemaglutinin	1eo8	B	A B	7.0E-01	9.9E-02	2.0E-01
Fab F9.13.7	1fbi	A	L H	5.5E-01	3.5E-01	1.0E-01
Lysozyme	1fbi	B	X	4.8E-01	2.6E-01	2.6E-01
Fab nmc-4	1fns	A	L H	5.8E-01	2.9E-01	1.2E-01
von Willebrand factor	1fns	B	A	5.0E-01	1.6E-01	3.4E-01
Kappa Fab	1fsk	A	K L	6.1E-01	3.1E-01	8.2E-02
major pollen antigen bet v 1-a	1fsk	B	J	5.2E-01	2.6E-01	2.2E-01
Fab 17b	1gc1	A	H L	6.0E-01	3.1E-01	8.3E-02
gp120	1gc1	B	G C	5.9E-01	2.1E-01	2.0E-01
Fab 730.1.4	1iai	A	L H	5.2E-01	4.5E-01	3.9E-02
Fab 409.5.3	1iai	B	M I	5.7E-01	4.0E-01	3.6E-02
Fab	1jhl	A	L H	5.2E-01	4.1E-01	6.9E-02
protein G domain III	1jhl	B	A	5.2E-01	2.9E-01	1.9E-01
Fab A6	1jrh	A	L H	5.3E-01	3.4E-01	1.2E-01
interferon gamma receptor alpha cha	1jrh	B	I	5.8E-01	3.0E-01	1.2E-01
Fab desire-1	1kb5	A	L H	5.0E-01	3.9E-01	1.1E-01
Kb5-c20 T-cell receptor	1kb5	B	A B	5.2E-01	2.7E-01	2.1E-01
Lysozyme	1mel	A	L	5.2E-01	2.5E-01	2.3E-01
VH single domain antibody	1mel	B	A	7.5E-01	2.4E-01	9.0E-03
Lysozyme	1mlc	A	E	5.8E-01	2.4E-01	1.8E-01
Fab D44.1	1mlc	B	A B	5.8E-01	3.8E-01	4.6E-02
Fab	1nca	A	L H	5.0E-01	4.0E-01	1.0E-01
N9 neuraminidase-nc41	1nca	B	N	5.7E-01	3.5E-01	8.1E-02
N15 α-β T-cell receptor	1nfd	A	A B	5.1E-01	3.3E-01	1.6E-01
Fab h57	1nfd	B	E F	5.3E-01	1.6E-01	3.1E-01
Fab NC-10	1nmb	A	L H	5.1E-01	4.1E-01	8.0E-02
N9-neuraminidase	1nmb	B	N	5.8E-01	3.9E-01	2.9E-02
Kappa Fab	1nsn	A	L H	5.3E-01	4.2E-01	5.0E-02
staphylococcal nuclease	1nsn	B	S	5.7E-01	2.0E-01	2.3E-01
Fab 184.1	1osp	A	L H	5.4E-01	2.0E-01	2.6E-01
protein A outer surface	1osp	B	O	5.4E-01	2.6E-01	2.0E-01
Fab	1qfw	A	HLIM	5.2E-01	3.6E-01	1.2E-01
Gonadotropin	1qfw	B	A B	5.8E-01	2.7E-01	1.5E-01
rhinovirus coat protein	1rvf	A	VWXY	4.8E-01	2.2E-01	3.1E-01
Fab 17-1a	1rvf	B	H L	5.4E-01	3.1E-01	1.5E-01
Fv D1.3	1vfb	A	A B	5.0E-01	3.0E-01	1.9E-01
Lysozyme	1vfb	B	C	4.8E-01	3.8E-01	1.4E-01
Fab e8	1wej	A	H L	5.0E-01	4.2E-01	8.2E-02
cytochrome C	1wej	B	F	5.3E-01	2.0E-01	2.7E-01
lambda Fab	2vir	A	A B	5.5E-01	3.3E-01	1.2E-01
hemagglutinin	2vir	B	C	5.2E-01	4.3E-01	4.8E-02
hyhel-5	3hfl	A	L H	5.2E-01	3.3E-01	1.5E-01
Lysozyme	3hfl	B	Y	5.1E-01	3.4E-01	1.5E-01
hyhel-10	3hfm	A	L H	4.4E-01	5.4E-01	2.0E-02
Lysozyme	3hfm	B	Y	5.3E-01	2.0E-01	2.7E-01

	pdbid			Resolution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Enzyme complexes								
Actin	1atn	A	A	2.6	41357	71260	16600	9.4E+02
DNAase I	1atn	B	D	2.6	28814	47770	10270	8.1E+02
Barley alpha-amylase/subtilisin inhibi	1ava	A	D	1.8	19846	34880	8645	1.2E+03
Barley alpha-amylase 2	1ava	B	B	1.8	44842	74310	14920	1.1E+03
VC1	1azs	A	A	2	42073	73540	17180	1.0E+03
IIC2	1azs	B	B	2	39502	69740	17140	8.9E+02
Cyclin-dependent kinase 6	1bi8	A	A	3.2	29889	52790	13460	1.0E+03
cyclin-dependent kinase inhibitor	1bi8	B	B	3.2	16392	28080	7085	9.4E+02
plasmin	1bml	A	A B	1.9	54486	93000	10590	2.1E+03
streptokinase	1bml	B	C	1.9	36079	66080	18040	2.0E+03
Barnase	1brs	A	E	2.2	12162	21690	5837	7.8E+02
Barstar C40A, C82A	1brs	B	I	2.2	9758	17880	4800	7.6E+02
plasmin	1bui	A	A B	3	52892	90060	20600	1.7E+03
staphylokinase sak-c-phi-c	1bui	B	C	3	13473	24940	7111	1.5E+03
m-calpain	1df0	A	A	2.3	70253	124500	30630	2.8E+03
calpain	1df0	B	B	2.3	20279	37210	10620	2.7E+03
ribonuclease A	1dfj	A	E	2.6	13667	24670	6921	1.3E+03
ribonuclease inhibitor	1dfj	B	I	2.6	48929	84370	18590	1.2E+03
porcine pancreatic α-amylase	1dhk	A	A	2.4	55145	89950	17760	1.4E+03
bean lectin-like inhibitor	1dhk	B	B	2.4	21517	37340	8871	1.6E+03
DNA polymerase processivity factor	1dml	A	A	1.6	29194	52340	13240	1.4E+03
DNA polymerase	1dml	B	B	1.6	3739	8036	3262	1.4E+03
triacylglycerol acyl-hydrolase	1eth	A	A	3.1	49768	81490	17860	7.4E+02
colipase	1eth	B	B	3.1	9477	17300	5413	8.0E+02
thioredoxin reductase	1f6m	A	E F	2.2	68822	117400	13450	9.2E+02
thioredoxin 1	1f6m	B	G	2.2	11629	20650	5454	8.7E+02
botulinum neurotoxin type B	1f83	A	A	2.55	48924	83940	19440	1.8E+03
synaptobrevin II	1f83	B	B C	2.55	4058	9858	4617	2.3E+03
acetylcholinesterase	1fss	A	A	3	59753	98120	19830	9.4E+02
fasciculin II	1fss	B	B	3	6748	12560	4034	1.0E+03
glycerol kinase	1gla	A	F	3	17008	31030	8242	6.1E+02
glucose specific factor III	1gla	B	G	3	53559	88730	17960	6.7E+02
serine/threonine phosphatase B2	1tco	A	A B	2.95	59700	100900	22800	9.5E+02
Fk506-binding protein	1tco	B	C	2.95	11711	20900	5678	9.2E+02
uracil-dna glycosylase	1udi	A	E	2.96	25716	44660	10400	1.0E+03
uracil-dna glycosylase inhibitor prote	1udi	B	I	2.96	9314	17280	5044	9.8E+02
beta-actin	2btf	A	A	2.35	41525	71670	16570	1.0E+03
profilin	2btf	B	B	2.35	14900	26520	6797	1.0E+03
cytochrome C peroxidase	2pcc	A	A	2.3	33413	56590	12940	5.6E+02
iso-I-cytochrome C	2pcc	B	B	2.3	12027	21820	6109	5.7E+02
Large complexes								
ribonuclease inhibitor	1a4y	A	A D	2.8	48904	163200	16180	1.3E+03
angiogenin	1a4y	B	B	2.8	14128	25040	6864	1.4E+03
allophycocyanin α−β	1b33	A	ACEBDF	2.5	102954	176100	41090	2.0E+03
phycobilisome	1b33	B	N	2.5	7732	15600	5245	2.1E+03
tp7 mab	1bgx	A	L H	1.9	46003	83030	19410	2.9E+03
Taq DNA polymerase	1bgx	B	T	1.9	92869	171600	39390	2.9E+03
Hemachromatosis protein hfe	1de4	A	A	2.5	43302	75140	18580	1.0E+03
transferrin receptor	1de4	B	B	2.5	141842	234500	22820	1.0E+03

	pdbid			Circularity	RMSD to plane	Atom burial	Atom access-ibility
Enzyme complexes							
Actin	1atn	A	A	3.5E-01	2.8E+00	1.1E+01	1.5E+01
DNAase I	1atn	B	D	6.0E-01	3.1E+00	7.5E+00	1.1E+01
Barley alpha-amylase/subtilisin inhibi	1ava	A	D	4.5E-01	3.6E+00	8.4E+00	1.3E+01
Barley alpha-amylase 2	1ava	B	B	3.6E-01	3.9E+00	8.4E+00	1.4E+01
VC1	1azs	A	A	8.5E-01	3.5E+00	1.0E+01	1.5E+01
IIC2	1azs	B	B	7.5E-01	3.3E+00	8.2E+00	1.3E+01
Cyclin-dependent kinase 6	1bi8	A	A	6.7E-01	2.6E+00	9.3E+00	1.4E+01
cyclin-dependent kinase inhibitor	1bi8	B	B	8.4E-01	2.3E+00	8.2E+00	1.3E+01
plasmin	1bml	A	A B	5.9E-01	4.4E+00	1.0E+01	1.6E+01
streptokinase	1bml	B	C	5.0E-01	5.2E+00	8.2E+00	1.4E+01
Barnase	1brs	A	E	4.0E-01	2.6E+00	8.2E+00	1.2E+01
Barstar C40A, C82A	1brs	B	I	6.2E-01	2.6E+00	8.5E+00	1.2E+01
plasmin	1bui	A	A B	7.6E-01	5.5E+00	8.7E+00	1.4E+01
staphylokinase sak-c-phi-c	1bui	B	C	3.8E-01	5.2E+00	8.7E+00	1.4E+01
m-calpain	1df0	A	A	6.9E-01	5.7E+00	1.0E+01	1.5E+01
calpain	1df0	B	B	9.2E-01	5.3E+00	9.5E+00	1.4E+01
ribonuclease A	1dfj	A	E	2.7E-01	5.1E+00	9.2E+00	1.6E+01
ribonuclease inhibitor	1dfj	B	I	3.6E-01	5.3E+00	7.7E+00	1.5E+01
porcine pancreatic α-amylase	1dhk	A	A	6.8E-01	5.5E+00	8.1E+00	1.2E+01
bean lectin-like inhibitor	1dhk	B	B	8.6E-01	5.2E+00	9.0E+00	1.3E+01
DNA polymerase processivity factor	1dml	A	A	4.5E-01	4.0E+00	9.6E+00	1.3E+01
DNA polymerase	1dml	B	B	5.2E-01	2.9E+00	1.0E+01	1.7E+01
triacylglycerol acyl-hydrolase	1eth	A	A	1.8E-01	2.5E+00	8.9E+00	1.5E+01
colipase	1eth	B	B	1.5E-01	2.7E+00	1.0E+01	1.6E+01
thioredoxin reductase	1f6m	A	E F	2.2E-01	3.2E+00	8.7E+00	1.5E+01
thioredoxin 1	1f6m	B	G	2.7E-01	3.4E+00	9.2E+00	1.5E+01
botulinum neurotoxin type B	1f83	A	A	1.6E-01	4.9E+00	6.9E+00	1.1E+01
synaptobrevin II	1f83	B	B C	7.1E-02	4.1E+00	1.2E+01	1.8E+01
acetylcholinesterase	1fss	A	A	6.6E-01	3.1E+00	7.6E+00	1.2E+01
fasciculin II	1fss	B	B	4.1E-01	2.8E+00	1.0E+01	1.6E+01
glycerol kinase	1gla	A	F	5.9E-01	1.9E+00	8.2E+00	1.4E+01
glucose specific factor III	1gla	B	G	5.1E-01	1.9E+00	1.0E+01	1.9E+01
serine/threonine phosphatase B2	1tco	A	A B	6.4E-01	3.9E+00	9.5E+00	1.7E+01
Fk506-binding protein	1tco	B	C	8.4E-01	3.7E+00	8.9E+00	1.7E+01
uracil-dna glycosylase	1udi	A	E	8.5E-01	3.0E+00	9.8E+00	1.6E+01
uracil-dna glycosylase inhibitor prote	1udi	B	I	7.9E-01	3.0E+00	8.9E+00	1.4E+01
beta-actin	2btf	A	A	2.5E-01	2.8E+00	8.1E+00	1.2E+01
profilin	2btf	B	B	4.6E-01	3.4E+00	7.3E+00	1.1E+01
cytochrome C peroxidase	2pcc	A	A	8.1E-01	1.4E+00	1.0E+01	1.8E+01
iso-I-cytochrome C	2pcc	B	B	7.5E-01	1.4E+00	1.0E+01	2.1E+01
Large complexes							
ribonuclease inhibitor	1a4y	A	A D	5.0E-01	5.9E+00	7.6E+00	1.3E+01
angiogenin	1a4y	B	B	3.4E-01	4.9E+00	9.8E+00	1.6E+01
allophycocyanin α-β	1b33	A	ACEBDF	2.9E-01	5.5E+00	8.2E+00	1.4E+01
phycobilisome	1b33	B	N	1.6E-01	4.7E+00	1.0E+01	1.7E+01
tp7 mab	1bgx	A	L H	5.4E-01	6.2E+00	8.6E+00	1.3E+01
Taq DNA polymerase	1bgx	B	T	5.0E-01	6.5E+00	9.4E+00	1.5E+01
Hemachromatosis protein hfe	1de4	A	A	7.6E-01	1.9E+00	8.8E+00	1.5E+01
transferrin receptor	1de4	B	B	5.0E-01	6.1E+00	9.1E+00	1.5E+01

	pdbid			Shape comple- mentarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Enzyme complexes						
Actin	1atn	A	A	7.4E-01	1.0E+00	1.8E+00
DNAase I	1atn	B	D	7.4E-01	1.0E+00	1.9E+00
Barley alpha-amylase/subtilisin inhibi	1ava	A	D	6.4E-01	1.1E+00	1.6E+00
Barley alpha-amylase 2	1ava	B	B	6.4E-01	9.9E-01	1.6E+00
VC1	1azs	A	A	6.6E-01	1.2E+00	1.8E+00
IIC2	1azs	B	B	6.6E-01	1.2E+00	2.2E+00
Cyclin-dependent kinase 6	1bi8	A	A	6.3E-01	1.1E+00	2.3E+00
cyclin-dependent kinase inhibitor	1bi8	B	B	6.3E-01	1.1E+00	2.0E+00
plasmin	1bml	A	A B	6.5E-01	9.9E-01	2.2E+00
streptokinase	1bml	B	C	6.5E-01	1.2E+00	2.0E+00
Barnase	1brs	A	E	7.2E-01	1.0E+00	1.3E+00
Barstar C40A, C82A	1brs	B	I	7.2E-01	9.2E-01	1.2E+00
plasmin	1bui	A	A B	6.0E-01	1.0E+00	1.9E+00
staphylokinase sak-c-phi-c	1bui	B	C	6.0E-01	1.1E+00	1.9E+00
m-calpain	1df0	A	A	6.5E-01	1.2E+00	2.1E+00
calpain	1df0	B	B	6.5E-01	1.1E+00	2.0E+00
ribonuclease A	1dfj	A	E	5.8E-01	1.2E+00	2.0E+00
ribonuclease inhibitor	1dfj	B	I	5.8E-01	1.1E+00	2.4E+00
porcine pancreatic α-amylase	1dhk	A	A	6.2E-01	1.0E+00	1.6E+00
bean lectin-like inhibitor	1dhk	B	B	6.2E-01	9.8E-01	1.6E+00
DNA polymerase processivity factor	1dml	A	A	7.1E-01	1.1E+00	2.0E+00
DNA polymerase	1dml	B	B	7.1E-01	1.4E+00	2.7E+00
triacylglycerol acyl-hydrolase	1eth	A	A	6.4E-01	1.1E+00	2.0E+00
colipase	1eth	B	B	6.4E-01	1.2E+00	2.2E+00
thioredoxin reductase	1f6m	A	E F	6.7E-01	1.1E+00	1.6E+00
thioredoxin 1	1f6m	B	G	6.7E-01	1.1E+00	1.5E+00
botulinum neurotoxin type B	1f83	A	A	6.1E-01	1.0E+00	1.5E+00
synaptobrevin II	1f83	B	B C	6.1E-01	9.6E-01	2.2E+00
acetylcholinesterase	1fss	A	A	6.9E-01	1.1E+00	2.0E+00
fasciculin II	1fss	B	B	6.9E-01	1.0E+00	2.0E+00
glycerol kinase	1gla	A	F	6.3E-01	1.1E+00	1.9E+00
glucose specific factor III	1gla	B	G	6.3E-01	1.1E+00	2.3E+00
serine/threonine phosphatase B2	1tco	A	A B	7.1E-01	9.5E-01	1.5E+00
Fk506-binding protein	1tco	B	C	7.1E-01	1.1E+00	1.6E+00
uracil-dna glycosylase	1udi	A	E	6.4E-01	1.2E+00	1.9E+00
uracil-dna glycosylase inhibitor prote	1udi	B	I	6.4E-01	1.5E+00	2.1E+00
beta-actin	2btf	A	A	6.9E-01	1.1E+00	2.1E+00
profilin	2btf	B	B	6.9E-01	1.1E+00	2.1E+00
cytochrome C peroxidase	2pcc	A	A	5.6E-01	8.3E-01	1.8E+00
iso-l-cytochrome C	2pcc	B	B	5.6E-01	1.0E+00	2.0E+00
Large complexes						
ribonuclease inhibitor	1a4y	A	A D	6.8E-01	1.0E+00	1.8E+00
angiogenin	1a4y	B	B	6.8E-01	1.3E+00	2.0E+00
allophycocyanin α−β	1b33	A	ACEBDF	7.0E-01	1.1E+00	1.7E+00
phycobilisome	1b33	B	N	7.0E-01	1.2E+00	1.9E+00
tp7 mab	1bgx	A	L H	5.1E-01	1.0E+00	1.8E+00
Taq DNA polymerase	1bgx	B	T	5.1E-01	1.0E+00	2.2E+00
Hemachromatosis protein hfe	1de4	A	A	7.3E-01	1.1E+00	2.0E+00
transferrin receptor	1de4	B	B	7.3E-01	1.0E+00	2.0E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
Enzyme complexes							
Actin	1atn	A	A	8.0E+00	8.5E-01	1.8E+01	1.9E+00
DNAase I	1atn	B	D	8.0E+00	9.8E-01	1.8E+01	2.2E+00
Barley alpha-amylase/subtilisin inhibi	1ava	A	D	8.0E+00	6.9E-01	1.9E+01	1.6E+00
Barley alpha-amylase 2	1ava	B	B	8.0E+00	7.0E-01	1.9E+01	1.7E+00
VC1	1azs	A	A	2.7E+01	2.7E+00	4.2E+01	4.2E+00
IIC2	1azs	B	B	2.7E+01	3.0E+00	4.2E+01	4.7E+00
Cyclin-dependent kinase 6	1bi8	A	A	4.0E+00	3.9E-01	1.5E+01	1.5E+00
cyclin-dependent kinase inhibitor	1bi8	B	B	4.0E+00	4.2E-01	1.5E+01	1.6E+00
plasmin	1bml	A	A B	2.8E+01	1.3E+00	4.4E+01	2.1E+00
streptokinase	1bml	B	C	2.8E+01	1.4E+00	4.4E+01	2.2E+00
Barnase	1brs	A	E	1.1E+01	1.4E+00	2.2E+01	2.8E+00
Barstar C40A, C82A	1brs	B	I	1.1E+01	1.4E+00	2.2E+01	2.9E+00
plasmin	1bui	A	A B	1.2E+01	7.2E-01	2.7E+01	1.6E+00
staphylokinase sak-c-phi-c	1bui	B	C	1.2E+01	7.8E-01	2.7E+01	1.8E+00
m-calpain	1df0	A	A	1.6E+01	5.8E-01	2.2E+01	8.0E-01
calpain	1df0	B	B	1.6E+01	5.9E-01	2.2E+01	8.2E-01
ribonuclease A	1dfj	A	E	1.5E+01	1.1E+00	2.0E+01	1.5E+00
ribonuclease inhibitor	1dfj	B	I	1.5E+01	1.2E+00	2.0E+01	1.6E+00
porcine pancreatic α-amylase	1dhk	A	A	8.0E+00	5.5E-01	1.5E+01	1.0E+00
bean lectin-like inhibitor	1dhk	B	B	8.0E+00	5.1E-01	1.5E+01	9.6E-01
DNA polymerase processivity factor	1dml	A	A	1.3E+01	9.5E-01	2.1E+01	1.5E+00
DNA polymerase	1dml	B	B	1.3E+01	9.4E-01	2.1E+01	1.5E+00
triacylglycerol acyl-hydrolase	1eth	A	A	2.6E+01	3.5E+00	5.0E+01	6.7E+00
colipase	1eth	B	B	2.6E+01	3.2E+00	5.0E+01	6.2E+00
thioredoxin reductase	1f6m	A	E F	1.4E+01	1.5E+00	2.7E+01	2.9E+00
thioredoxin 1	1f6m	B	G	1.4E+01	1.6E+00	2.7E+01	3.1E+00
botulinum neurotoxin type B	1f83	A	A	1.5E+01	8.4E-01	2.7E+01	1.5E+00
synaptobrevin II	1f83	B	B C	1.5E+01	6.5E-01	2.7E+01	1.2E+00
acetylcholinesterase	1fss	A	A	8.0E+00	8.5E-01	1.1E+01	1.2E+00
fasciculin II	1fss	B	B	8.0E+00	7.9E-01	1.1E+01	1.1E+00
glycerol kinase	1gla	A	F	1.4E+01	2.3E+00	2.8E+01	4.6E+00
glucose specific factor III	1gla	B	G	1.4E+01	2.1E+00	2.8E+01	4.2E+00
serine/threonine phosphatase B2	1tco	A	A B	8.0E+00	8.4E-01	1.0E+01	1.1E+00
Fk506-binding protein	1tco	B	C	8.0E+00	8.7E-01	1.0E+01	1.1E+00
uracil-dna glycosylase	1udi	A	E	2.0E+01	2.0E+00	4.1E+01	4.0E+00
uracil-dna glycosylase inhibitor prote	1udi	B	I	2.0E+01	2.0E+00	4.1E+01	4.2E+00
beta-actin	2btf	A	A	7.0E+01	6.8E+00	1.2E+02	1.1E+01
profilin	2btf	B	B	7.0E+01	6.9E+00	1.2E+02	1.2E+01
cytochrome C peroxidase	2pcc	A	A	3.0E+01	5.4E+00	4.8E+01	8.6E+00
iso-I-cytochrome C	2pcc	B	B	3.0E+01	5.3E+00	4.8E+01	8.4E+00
Large complexes							
ribonuclease inhibitor	1a4y	A	A D	1.5E+01	1.2E+00	2.3E+01	1.8E+00
angiogenin	1a4y	B	B	1.5E+01	1.1E+00	2.3E+01	1.7E+00
allophycocyanin α−β	1b33	A	ACEBDF	1.2E+01	5.9E-01	2.5E+01	1.2E+00
phycobilisome	1b33	B	N	1.2E+01	5.7E-01	2.5E+01	1.2E+00
tp7 mab	1bgx	A	L H	1.3E+01	4.5E-01	2.3E+01	8.0E-01
Taq DNA polymerase	1bgx	B	T	1.3E+01	4.5E-01	2.3E+01	8.0E-01
Hemachromatosis protein hfe	1de4	A	A	1.3E+01	1.3E+00	3.2E+01	3.2E+00
transferrin receptor	1de4	B	B	1.3E+01	1.2E+00	3.2E+01	3.1E+00

	pdbid			%non-pola	%polar	%charged
Enzyme complexes						
Actin	1atn	A	A	7.1E-01	2.8E-01	1.3E-02
DNAase I	1atn	B	D	7.4E-01	2.4E-01	1.8E-02
Barley alpha-amylase/subtilisin inhibi	1ava	A	D	5.5E-01	2.8E-01	1.7E-01
Barley alpha-amylase 2	1ava	B	B	6.1E-01	2.9E-01	9.9E-02
VC1	1azs	A	A	7.1E-01	2.0E-01	9.4E-02
IIC2	1azs	B	B	5.5E-01	2.9E-01	1.6E-01
Cyclin-dependent kinase 6	1bi8	A	A	5.6E-01	1.6E-01	2.8E-01
cyclin-dependent kinase inhibitor	1bi8	B	B	6.2E-01	1.7E-01	2.1E-01
plasmin	1bml	A	A B	5.6E-01	2.5E-01	2.0E-01
streptokinase	1bml	B	C	5.8E-01	2.4E-01	1.7E-01
Barnase	1brs	A	E	5.5E-01	1.8E-01	2.7E-01
Barstar C40A, C82A	1brs	B	I	5.6E-01	2.4E-01	2.0E-01
plasmin	1bui	A	A B	5.8E-01	2.4E-01	1.8E-01
staphylokinase sak-c-phi-c	1bui	B	C	6.5E-01	2.8E-01	7.7E-02
m-calpain	1df0	A	A	7.1E-01	1.2E-01	1.8E-01
calpain	1df0	B	B	5.9E-01	2.5E-01	1.6E-01
ribonuclease A	1dfj	A	E	5.2E-01	3.0E-01	1.8E-01
ribonuclease inhibitor	1dfj	B	I	4.7E-01	2.8E-01	2.5E-01
porcine pancreatic α-amylase	1dhk	A	A	6.1E-01	2.4E-01	1.5E-01
bean lectin-like inhibitor	1dhk	B	B	5.8E-01	3.2E-01	9.9E-02
DNA polymerase processivity factor	1dml	A	A	6.9E-01	1.8E-01	1.3E-01
DNA polymerase	1dml	B	B	6.9E-01	1.9E-01	1.2E-01
triacylglycerol acyl-hydrolase	1eth	A	A	5.4E-01	2.7E-01	1.8E-01
colipase	1eth	B	B	4.9E-01	1.4E-01	3.8E-01
thioredoxin reductase	1f6m	A	E F	5.6E-01	3.4E-01	1.0E-01
thioredoxin 1	1f6m	B	G	6.0E-01	2.5E-01	1.5E-01
botulinum neurotoxin type B	1f83	A	A	5.6E-01	2.1E-01	2.3E-01
synaptobrevin II	1f83	B	B C	5.7E-01	2.5E-01	1.9E-01
acetylcholinesterase	1fss	A	A	5.9E-01	2.9E-01	1.2E-01
fasciculin II	1fss	B	B	5.7E-01	2.5E-01	1.9E-01
glycerol kinase	1gla	A	F	7.4E-01	7.5E-02	1.8E-01
glucose specific factor III	1gla	B	G	5.5E-01	2.1E-01	2.4E-01
serine/threonine phosphatase B2	1tco	A	A B	5.4E-01	3.6E-01	1.0E-01
Fk506-binding protein	1tco	B	C	6.0E-01	2.2E-01	1.8E-01
uracil-dna glycosylase	1udi	A	E	6.7E-01	2.4E-01	9.4E-02
uracil-dna glycosylase inhibitor prote	1udi	B	I	5.9E-01	2.9E-01	1.2E-01
beta-actin	2btf	A	A	5.6E-01	1.9E-01	2.5E-01
profilin	2btf	B	B	6.1E-01	2.1E-01	1.9E-01
cytochrome C peroxidase	2pcc	A	A	4.7E-01	3.3E-01	2.0E-01
iso-I-cytochrome C	2pcc	B	B	6.0E-01	1.9E-01	2.2E-01
Large complexes						
ribonuclease inhibitor	1a4y	A	A D	5.0E-01	3.1E-01	2.0E-01
angiogenin	1a4y	B	B	5.1E-01	2.1E-01	2.8E-01
allophycocyanin α–β	1b33	A	ACEBDF	4.6E-01	3.1E-01	2.2E-01
phycobilisome	1b33	B	N	5.7E-01	3.4E-01	9.0E-02
tp7 mab	1bgx	A	L H	5.8E-01	3.4E-01	7.8E-02
Taq DNA polymerase	1bgx	B	T	6.3E-01	2.0E-01	1.7E-01
Hemachromatosis protein hfe	1de4	A	A	6.7E-01	2.5E-01	8.3E-02
transferrin receptor	1de4	B	B	5.9E-01	2.0E-01	2.1E-01

	pdbid			Resolution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Large complexes								
dihydrolipoamide dehydrogenase	1ebd	A	A B	1.95	95042	162100	34620	6.1E+02
dihydrolipoamide acetyltransferase	1ebd	B	C	1.95	4320	8691	2992	6.4E+02
Fe-S and Mo subunits carbon mono:	1ffu	A	ADBE	2.1	206155	334800	61200	4.0E+03
Flavoprotein carbon monoxide	1ffu	B	CF	2.1	60666	106300	53180	4.6E+03
Fe-S and flavo subunits carbon mon	1ffu	A	ADCF	2.4	94693	384000	25290	6.5E+03
Mo subunit carbon monoxide	1ffu	B	BE	2.4	172128	61740	74320	6.7E+03
Mo and flavo subunits carbon mono:	1ffu	A	BECF	2.65	232794	163500	36900	6.1E+03
Fe-S subunit carbon monoxide	1ffu	B	AD	2.65	34027	281300	16730	6.1E+03
nitrogenase Mo-Fe	1n2c	A	ABCD	1.9	225968	353200	57690	3.7E+03
nitrogenase Fe	1n2c	B	EF	1.9	59180	97380	18270	1.8E+03
Cytochrome C oxidase	1qle	A	ABCD	3	123414	206700	41400	6.4E+02
Fab	1qle	B	HL	3	25077	42800	10220	6.8E+02
HIV reverse transcriptase	2hmi	A	A B	1.6	107412	189900	45060	6.2E+02
Fab 28	2hmi	B	C D	1.6	46330	79700	19160	6.0E+02
G-protein, cell cycle, signal transduction								
C-fos, C-jun	1a02	A	F J	2.5	12430	25650	9342	6.9E+02
NFAT	1a02	B	N	2.5	31239	56910	15260	7.0E+02
CheY	1a0o	A	A	2	13933	24940	6472	5.7E+02
CheA	1a0o	B	B	2	7438	14200	4213	5.5E+02
nuclear transport factor 2	1a2k	A	A B	3.15	28202	49020	11620	7.3E+02
ran	1a2k	B	C	3.15	22162	39310	9764	8.6E+02
guanine nucleotide-binding protein (	1agr	A	A	2.6	39794	68870	17420	8.2E+02
RGS4	1agr	B	E	2.6	15060	26590	7428	8.0E+02
elongation factor TU T. thermophilus	1aip	A	E	2.8	40695	71400	17380	1.4E+03
elongation factor TS T. thermophilus	1aip	B	G H	2.8	43671	78510	21090	2.9E+03
Fk506-binding protein	1b6c	A	A	1.85	11781	21320	5924	8.5E+02
TGF-B superfamily receptor type 1	1b6c	B	B	1.85	36974	64130	15270	8.8E+02
death domain of pelle	1d2z	A	A	2	11772	21600	5975	7.8E+02
death domain of tube	1d2z	B	B	2	16661	30360	8407	8.9E+02
mothers agains decapentaplegic	1dev	A	A	2.9	21998	37420	9223	1.2E+03
smad anchor for receptor activation	1dev	B	B	2.9	4125	9339	4203	1.4E+03
fyn tyrosine kinase	1efn	A	A	2.7	5990	11130	3486	6.2E+02
HIV-1 NEF protein	1efn	B	B	2.7	12304	22030	6169	6.2E+02
elongation factor TU E. coli	1efu	A	A	2.37	39726	69060	16620	1.7E+03
elongation factor TS E. coli	1efu	B	B	2.37	30218	54980	15480	1.9E+03
dimerization cofactor of hepatocyte r	1f93	A	A B	3.2	22476	40600	10850	9.4E+02
hepatocyte nuclear factor 1-a	1f93	B	E F	3.2	6017	12200	4099	9.6E+02
cbl	1fbv	A	A	2.3	45444	78510	19040	8.7E+02
zap-70	1fbv	B	B	2.3	16280	30740	8627	9.0E+02
cyclin-dependent kinase-2	1fin	A	A	2.3	33871	58500	13930	1.6E+03
cyclin-A	1fin	B	B	2.3	29773	51150	12200	1.8E+03
ralgds	1lfd	A	A C	2.9	19868	36390	10120	1.8E+03
ras	1lfd	B	B D	2.9	37852	65640	15750	1.8E+03
cyclin-dependent kinase-6 and inhib	1g3n	A	A E	2	99786	115700	25530	2.4E+03
v-cyclin	1g3n	B	G	2	51102	44900	24950	1.2E+03
cyclin-dependent kinase-6 and V-cyc	1g3n	A	A E	2.9	117376	115700	10750	1.8E+03
cyclin-dependent kinase-6 inhibitor	1g3n	B	B	2.9	33512	29200	7360	8.2E+02
Giα1	1gg2	A	A	3.5	39221	69110	17760	1.2E+03
Giβ1	1gg2	B	B G	3.5	43094	71800	16360	1.1E+03

	pdbid			Circularity	RMSD to plane	Atom burial	Atom access-ibility
Large complexes							
dihydrolipoamide dehydrogenase	1ebd	A	A B	3.5E-01	2.6E+00	8.0E+00	1.5E+01
dihydrolipoamide acetyltransferase	1ebd	B	C	2.1E-01	2.3E+00	1.2E+01	1.8E+01
Fe-S and Mo subunits carbon mono:	1ffu	A	ADBE	5.8E-02	8.2E+00	8.0E+00	1.3E+01
Flavoprotein carbon monoxide	1ffu	B	CF	4.5E-02	7.8E+00	1.1E+01	1.4E+01
Fe-S and flavo subunits carbon mon	1ffu	A	ADCF	2.0E-01	8.2E+00	9.1E+00	1.5E+01
Mo subunit carbon monoxide	1ffu	B	BE	1.5E-01	7.4E+00	9.3E+00	1.2E+01
Mo and flavo subunits carbon mono:	1ffu	A	BECF	8.5E-02	7.9E+00	8.9E+00	1.3E+01
Fe-S subunit carbon monoxide	1ffu	B	AD	8.1E-02	8.1E+00	1.0E+01	1.3E+01
nitrogenase Mo-Fe	1n2c	A	ABCD	5.3E-02	9.5E+00	9.5E+00	1.5E+01
nitrogenase Fe	1n2c	B	EF	8.2E-01	3.9E+00	8.4E+00	1.3E+01
Cytochrome C oxidase	1qle	A	ABCD	1.7E-01	1.5E+00	8.2E+00	1.3E+01
Fab	1qle	B	HL	2.5E-01	1.4E+00	9.1E+00	1.5E+01
HIV reverse transcriptase	2hmi	A	A B	4.5E-01	2.7E+00	8.1E+00	1.4E+01
Fab 28	2hmi	B	C D	7.1E-01	2.4E+00	7.7E+00	1.3E+01
G-protein, cell cycle, signal transduction							
C-fos, C-jun	1a02	A	F J	2.5E-01	3.3E+00	8.6E+00	1.9E+01
NFAT	1a02	B	N	4.5E-01	3.8E+00	9.4E+00	2.2E+01
CheY	1a0o	A	A	3.2E-01	2.7E+00	1.0E+01	1.7E+01
CheA	1a0o	B	B	3.8E-01	2.6E+00	8.2E+00	1.5E+01
nuclear transport factor 2	1a2k	A	A B	4.1E-01	2.7E+00	8.9E+00	1.4E+01
ran	1a2k	B	C	4.8E-01	2.7E+00	1.1E+01	1.5E+01
guanine nucleotide-binding protein (	1agr	A	A	8.2E-01	2.3E+00	9.0E+00	1.4E+01
RGS4	1agr	B	E	6.1E-01	2.2E+00	7.8E+00	1.3E+01
elongation factor TU T. thermophilus	1aip	A	E	7.6E-02	3.1E+00	9.0E+00	1.4E+01
elongation factor TS T. thermophilus	1aip	B	G H	1.5E-01	5.4E+00	9.9E+00	1.6E+01
Fk506-binding protein	1b6c	A	A	8.2E-01	2.9E+00	8.0E+00	1.2E+01
TGF-B superfamily receptor type 1	1b6c	B	B	7.6E-01	2.5E+00	9.4E+00	1.4E+01
death domain of pelle	1d2z	A	A	3.9E-01	2.8E+00	7.4E+00	1.3E+01
death domain of tube	1d2z	B	B	2.1E-01	3.1E+00	1.1E+01	1.7E+01
mothers agains decapentaplegic	1dev	A	A	2.3E-01	3.4E+00	7.0E+00	1.3E+01
smad anchor for receptor activation	1dev	B	B	2.7E-01	2.8E+00	1.0E+01	1.8E+01
fyn tyrosine kinase	1efn	A	A	5.1E-01	1.8E+00	9.0E+00	1.4E+01
HIV-1 NEF protein	1efn	B	B	3.6E-01	1.8E+00	9.4E+00	1.4E+01
elongation factor TU E. coli	1efu	A	A	3.5E-01	3.3E+00	8.5E+00	1.4E+01
elongation factor TS E. coli	1efu	B	B	4.1E-01	3.3E+00	1.0E+01	1.6E+01
dimerization cofactor of hepatocyte r	1f93	A	A B	5.6E-01	2.1E+00	9.6E+00	1.4E+01
hepatocyte nuclear factor 1-a	1f93	B	E F	6.3E-01	2.0E+00	1.1E+01	1.6E+01
cbl	1fbv	A	A	1.7E-01	3.6E+00	8.5E+00	1.4E+01
zap-70	1fbv	B	B	2.2E-01	3.4E+00	9.9E+00	1.6E+01
cyclin-dependent kinase-2	1fin	A	A	2.8E-01	3.4E+00	8.2E+00	1.3E+01
cyclin-A	1fin	B	B	2.7E-01	3.4E+00	1.0E+01	1.4E+01
ralgds	1lfd	A	A C	2.9E-01	5.3E+00	9.8E+00	1.5E+01
ras	1lfd	B	B D	2.4E-01	5.1E+00	9.4E+00	1.5E+01
cyclin-dependent kinase-6 and inhibi	1g3n	A	A E	3.9E-01	3.0E+00	9.6E+00	1.5E+01
v-cyclin	1g3n	B	G	2.4E-01	2.3E+00	9.2E+00	1.4E+01
cyclin-dependent kinase-6 and V-cyc	1g3n	A	A E	5.2E-02	4.4E+00	9.8E+00	1.5E+01
cyclin-dependent kinase-6 inhibitor	1g3n	B	B	7.0E-01	2.2E+00	8.4E+00	1.2E+01
Giα1	1gg2	A	A	1.6E-01	5.2E+00	9.6E+00	1.7E+01

	pdbid			Shape comple-mentarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Large complexes						
dihydrolipoamide dehydrogenase	1ebd	A	A B	6.9E-01	8.8E-01	2.3E+00
dihydrolipoamide acetyltransferase	1ebd	B	C	6.9E-01	8.2E-01	2.4E+00
Fe-S and Mo subunits carbon mono:	1ffu	A	ADBE	7.0E-01	1.0E+00	2.1E+00
Flavoprotein carbon monoxide	1ffu	B	CF	7.0E-01	1.2E+00	2.1E+00
Fe-S and flavo subunits carbon mon	1ffu	A	ADCF	7.1E-01	1.2E+00	2.3E+00
Mo subunit carbon monoxide	1ffu	B	BE	7.1E-01	1.1E+00	2.3E+00
Mo and flavo subunits carbon mono:	1ffu	A	BECF	7.3E-01	1.2E+00	2.4E+00
Fe-S subunit carbon monoxide	1ffu	B	AD	7.3E-01	1.2E+00	2.3E+00
nitrogenase Mo-Fe	1n2c	A	ABCD	7.0E-01	1.1E+00	2.2E+00
nitrogenase Fe	1n2c	B	EF	7.0E-01	1.0E+00	2.0E+00
Cytochrome C oxidase	1qle	A	ABCD	7.0E-01	1.0E+00	2.0E+00
Fab	1qle	B	HL	7.0E-01	1.1E+00	1.9E+00
HIV reverse transcriptase	2hmi	A	A B	7.1E-01	1.0E+00	1.9E+00
Fab 28	2hmi	B	C D	7.1E-01	9.9E-01	2.0E+00
G-protein, cell cycle, signal transduction						
C-fos, C-jun	1a02	A	F J	7.1E-01	1.1E+00	2.7E+00
NFAT	1a02	B	N	7.1E-01	8.9E-01	2.1E+00
CheY	1a0o	A	A	6.7E-01	1.1E+00	2.2E+00
CheA	1a0o	B	B	6.7E-01	1.0E+00	1.8E+00
nuclear transport factor 2	1a2k	A	A B	6.3E-01	1.3E+00	2.2E+00
ran	1a2k	B	C	6.3E-01	1.0E+00	1.6E+00
guanine nucleotide-binding protein (	1agr	A	A	7.3E-01	1.1E+00	2.0E+00
RGS4	1agr	B	E	7.3E-01	1.0E+00	2.1E+00
elongation factor TU T. thermophilus	1aip	A	E	6.5E-01	1.1E+00	2.3E+00
elongation factor TS T. thermophilus	1aip	B	G H	6.5E-01	1.2E+00	2.5E+00
Fk506-binding protein	1b6c	A	A	7.1E-01	1.1E+00	1.9E+00
TGF-B superfamily receptor type 1	1b6c	B	B	7.1E-01	1.0E+00	2.0E+00
death domain of pelle	1d2z	A	A	7.3E-01	1.1E+00	1.6E+00
death domain of tube	1d2z	B	B	7.3E-01	1.0E+00	1.6E+00
mothers agains decapentaplegic	1dev	A	A	8.0E-01	1.2E+00	1.9E+00
smad anchor for receptor activation	1dev	B	B	8.0E-01	1.4E+00	2.6E+00
fyn tyrosine kinase	1efn	A	A	7.7E-01	1.1E+00	1.8E+00
HIV-1 NEF protein	1efn	B	B	7.7E-01	1.1E+00	1.8E+00
elongation factor TU E. coli	1efu	A	A	7.7E-01	1.1E+00	1.6E+00
elongation factor TS E. coli	1efu	B	B	7.7E-01	1.2E+00	1.8E+00
dimerization cofactor of hepatocyte r	1f93	A	A B	7.1E-01	1.2E+00	1.5E+00
hepatocyte nuclear factor 1-a	1f93	B	E F	7.1E-01	9.5E-01	1.5E+00
cbl	1fbv	A	A	7.6E-01	1.1E+00	1.7E+00
zap-70	1fbv	B	B	7.6E-01	1.1E+00	1.9E+00
cyclin-dependent kinase-2	1fin	A	A	6.7E-01	1.0E+00	1.9E+00
cyclin-A	1fin	B	B	6.7E-01	1.2E+00	2.1E+00
ralgds	1lfd	A	A C	7.0E-01	1.1E+00	1.9E+00
ras	1lfd	B	B D	7.0E-01	1.0E+00	1.8E+00
cyclin-dependent kinase-6 and inhibi	1g3n	A	A E	6.6E-01	1.2E+00	1.6E+00
v-cyclin	1g3n	B	G	6.6E-01	1.3E+00	1.9E+00
cyclin-dependent kinase-6 and V-cyc	1g3n	A	A E	6.8E-01	1.1E+00	2.3E+00
cyclin-dependent kinase-6 inhibitor	1g3n	B	B	6.8E-01	1.0E+00	1.7E+00
Giα1	1gg2	A	A	8.0E-01	1.1E+00	1.9E+00
Giβ1	1gg2	B	B G	8.0E-01	1.3E+00	1.8E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
Large complexes							
dihydrolipoamide dehydrogenase	1ebd	A	A B	8.0E+00	1.3E+00	2.1E+01	3.4E+00
dihydrolipoamide acetyltransferase	1ebd	B	C	8.0E+00	1.2E+00	2.1E+01	3.3E+00
Fe-S and Mo subunits carbon mono:	1ffu	A	ADBE	9.0E+00	2.2E-01	1.9E+01	4.7E-01
Flavoprotein carbon monoxide	1ffu	B	CF	9.0E+00	2.0E-01	1.9E+01	4.2E-01
Fe-S and flavo subunits carbon mon	1ffu	A	ADCF	3.5E+01	5.4E-01	5.7E+01	8.8E-01
Mo subunit carbon monoxide	1ffu	B	BE	3.5E+01	5.2E-01	5.7E+01	8.5E-01
Mo and flavo subunits carbon mono:	1ffu	A	BECF	1.7E+01	2.8E-01	2.4E+01	3.9E-01
Fe-S subunit carbon monoxide	1ffu	B	AD	1.7E+01	2.8E-01	2.4E+01	4.0E-01
nitrogenase Mo-Fe	1n2c	A	ABCD	1.4E+01	3.8E-01	2.5E+01	6.8E-01
nitrogenase Fe	1n2c	B	EF	1.4E+01	7.8E-01	2.5E+01	1.4E+00
Cytochrome C oxidase	1qle	A	ABCD	1.0E+01	1.6E+00	2.4E+01	3.7E+00
Fab	1qle	B	HL	1.0E+01	1.5E+00	2.4E+01	3.5E+00
HIV reverse transcriptase	2hmi	A	A B	5.0E+01	8.0E+00	8.4E+01	1.3E+01
Fab 28	2hmi	B	C D	5.0E+01	8.4E+00	8.4E+01	1.4E+01
G-protein, cell cycle, signal transduction							
C-fos, C-jun	1a02	A	F J	1.0E+01	1.4E+00	1.5E+01	2.2E+00
NFAT	1a02	B	N	1.0E+01	1.4E+00	1.5E+01	2.1E+00
CheY	1a0o	A	A	1.0E+01	1.7E+00	1.6E+01	2.8E+00
CheA	1a0o	B	B	1.0E+01	1.8E+00	1.6E+01	2.9E+00
nuclear transport factor 2	1a2k	A	A B	9.0E+00	1.2E+00	1.4E+01	1.9E+00
ran	1a2k	B	C	9.0E+00	1.0E+00	1.4E+01	1.6E+00
guanine nucleotide-binding protein (	1agr	A	A	1.6E+01	1.9E+00	2.5E+01	3.0E+00
RGS4	1agr	B	E	1.6E+01	2.0E+00	2.5E+01	3.1E+00
elongation factor TU T. thermophilus	1aip	A	E	1.2E+01	8.7E-01	2.1E+01	1.5E+00
elongation factor TS T. thermophilus	1aip	B	G H	1.2E+01	4.1E-01	2.1E+01	7.2E-01
Fk506-binding protein	1b6c	A	A	1.6E+01	1.9E+00	2.8E+01	3.3E+00
TGF-B superfamily receptor type 1	1b6c	B	B	1.6E+01	1.8E+00	2.8E+01	3.2E+00
death domain of pelle	1d2z	A	A	1.4E+01	1.8E+00	2.7E+01	3.5E+00
death domain of tube	1d2z	B	B	1.4E+01	1.6E+00	2.7E+01	3.0E+00
mothers agains decapentaplegic	1dev	A	A	8.0E+00	6.8E-01	1.8E+01	1.5E+00
smad anchor for receptor activation	1dev	B	B	8.0E+00	5.7E-01	1.8E+01	1.3E+00
fyn tyrosine kinase	1efn	A	A	9.0E+00	1.5E+00	1.4E+01	2.3E+00
HIV-1 NEF protein	1efn	B	B	9.0E+00	1.4E+00	1.4E+01	2.3E+00
elongation factor TU E. coli	1efu	A	A	7.0E+00	4.1E-01	1.3E+01	7.7E-01
elongation factor TS E. coli	1efu	B	B	7.0E+00	3.7E-01	1.3E+01	6.9E-01
dimerization cofactor of hepatocyte r	1f93	A	A B	7.0E+00	7.5E-01	2.2E+01	2.3E+00
hepatocyte nuclear factor 1-a	1f93	B	E F	7.0E+00	7.3E-01	2.2E+01	2.3E+00
cbl	1fbv	A	A	3.0E+00	3.5E-01	7.0E+00	8.1E-01
zap-70	1fbv	B	B	3.0E+00	3.3E-01	7.0E+00	7.8E-01
cyclin-dependent kinase-2	1fin	A	A	1.9E+01	1.2E+00	3.6E+01	2.2E+00
cyclin-A	1fin	B	B	1.9E+01	1.1E+00	3.6E+01	2.0E+00
ralgds	1lfd	A	A C	1.3E+01	7.4E-01	2.2E+01	1.3E+00
ras	1lfd	B	B D	1.3E+01	7.1E-01	2.2E+01	1.2E+00
cyclin-dependent kinase-6 and inhibi	1g3n	A	A E	1.2E+01	5.0E-01	2.0E+01	8.4E-01
v-cyclin	1g3n	B	G	1.2E+01	1.0E+00	2.0E+01	1.7E+00
cyclin-dependent kinase-6 and V-cyc	1g3n	A	A E	1.5E+01	8.3E-01	2.7E+01	1.5E+00
cyclin-dependent kinase-6 inhibitor	1g3n	B	B	1.5E+01	1.8E+00	2.7E+01	3.3E+00
Giα1	1gg2	A	A	1.2E+01	1.0E+00	3.1E+01	2.6E+00
Giβ1	1gg2	B	B G	1.2E+01	1.1E+00	3.1E+01	2.8E+00

	pdbid			%non-pola	%polar	%charged
Large complexes						
dihydrolipoamide dehydrogenase	1ebd	A	A B	4.8E-01	2.1E-01	3.1E-01
dihydrolipoamide acetyltransferase	1ebd	B	C	5.6E-01	1.8E-01	2.7E-01
Fe-S and Mo subunits carbon mono:	1ffu	A	ADBE	6.3E-01	2.0E-01	1.7E-01
Flavoprotein carbon monoxide	1ffu	B	CF	6.2E-01	1.6E-01	2.2E-01
Fe-S and flavo subunits carbon mon	1ffu	A	ADCF	5.7E-01	2.3E-01	2.1E-01
Mo subunit carbon monoxide	1ffu	B	BE	6.7E-01	1.9E-01	1.4E-01
Mo and flavo subunits carbon mono:	1ffu	A	BECF	6.6E-01	2.1E-01	1.3E-01
Fe-S subunit carbon monoxide	1ffu	B	AD	5.8E-01	2.6E-01	1.7E-01
nitrogenase Mo-Fe	1n2c	A	ABCD	5.8E-01	2.3E-01	1.9E-01
nitrogenase Fe	1n2c	B	EF	6.1E-01	1.6E-01	2.3E-01
Cytochrome C oxidase	1qle	A	ABCD	5.6E-01	2.2E-01	2.2E-01
Fab	1qle	B	HL	3.9E-01	4.5E-01	1.6E-01
HIV reverse transcriptase	2hmi	A	A B	7.0E-01	1.1E-01	1.9E-01
Fab 28	2hmi	B	C D	6.4E-01	2.5E-01	1.1E-01
G-protein, cell cycle, signal transduction						
C-fos, C-jun	1a02	A	F J	3.5E-01	1.9E-01	4.6E-01
NFAT	1a02	B	N	5.1E-01	3.0E-01	1.9E-01
CheY	1a0o	A	A	6.1E-01	1.9E-01	2.0E-01
CheA	1a0o	B	B	6.3E-01	2.0E-01	1.6E-01
nuclear transport factor 2	1a2k	A	A B	6.5E-01	2.6E-01	9.8E-02
ran	1a2k	B	C	6.2E-01	2.2E-01	1.6E-01
guanine nucleotide-binding protein (	1agr	A	A	6.1E-01	2.2E-01	1.6E-01
RGS4	1agr	B	E	4.8E-01	2.6E-01	2.5E-01
elongation factor TU T. thermophilus	1aip	A	E	6.1E-01	2.4E-01	1.5E-01
elongation factor TS T. thermophilus	1aip	B	G H	5.9E-01	2.5E-01	1.6E-01
Fk506-binding protein	1b6c	A	A	6.7E-01	2.4E-01	8.8E-02
TGF-B superfamily receptor type 1	1b6c	B	B	7.1E-01	1.8E-01	1.1E-01
death domain of pelle	1d2z	A	A	5.8E-01	3.4E-01	8.0E-02
death domain of tube	1d2z	B	B	6.0E-01	2.6E-01	1.4E-01
mothers agains decapentaplegic	1dev	A	A	6.4E-01	3.1E-01	4.7E-02
smad anchor for receptor activation	1dev	B	B	7.5E-01	2.5E-01	3.2E-04
fyn tyrosine kinase	1efn	A	A	6.4E-01	2.8E-01	7.6E-02
HIV-1 NEF protein	1efn	B	B	6.3E-01	1.3E-01	2.5E-01
elongation factor TU E. coli	1efu	A	A	6.6E-01	2.2E-01	1.2E-01
elongation factor TS E. coli	1efu	B	B	6.5E-01	1.9E-01	1.6E-01
dimerization cofactor of hepatocyte r	1f93	A	A B	5.6E-01	7.1E-02	3.7E-01
hepatocyte nuclear factor 1-a	1f93	B	E F	7.1E-01	1.3E-01	1.6E-01
cbl	1fbv	A	A	6.3E-01	2.0E-01	1.7E-01
zap-70	1fbv	B	B	6.0E-01	8.3E-02	3.2E-01
cyclin-dependent kinase-2	1fin	A	A	6.3E-01	2.3E-01	1.4E-01
cyclin-A	1fin	B	B	6.4E-01	2.6E-01	1.1E-01
ralgds	1lfd	A	A C	5.5E-01	3.0E-01	1.5E-01
ras	1lfd	B	B D	5.2E-01	2.4E-01	2.4E-01
cyclin-dependent kinase-6 and inhibi	1g3n	A	A E	6.6E-01	1.9E-01	1.5E-01
v-cyclin	1g3n	B	G	6.6E-01	1.5E-01	1.9E-01
cyclin-dependent kinase-6 and V-cyc	1g3n	A	A E	5.6E-01	2.2E-01	2.2E-01
cyclin-dependent kinase-6 inhibitor	1g3n	B	B	6.6E-01	1.7E-01	1.8E-01
Giα1	1gg2	A	A	6.2E-01	1.8E-01	2.0E-01
Giβ1	1gg2	B	B G	5.5E-01	2.8E-01	1.7E-01

	pdbid			Reso-lution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
G-protein, cell cycle, signal transduction								
Giγ2	1got	A	A	4	37189	67710	18020	1.2E+03
Gtα-Gia	1got	B	B G	4	43975	72840	16100	1.2E+03
Gtβ	1gua	A	A	1.8	18983	33890	8654	6.4E+02
Gtγ	1gua	B	B	1.8	8673	16670	4921	6.3E+02
mcm1 transcriptiona regulator	1mnm	A	A B	2.3	18474	34890	9476	6.2E+02
mat-α-2 repressor	1mnm	B	C	2.3	8619	17200	5938	1.2E+03
ran	1rrp	A	A	2.7	23109	42250	11930	2.4E+03
nuclear core complex protein nup35	1rrp	B	B	2.7	15641	29430	8980	2.3E+03
p50-rhogap	1tx4	A	A	2.25	22257	39150	9593	1.1E+03
rhoa	1tx4	B	B	2.25	19621	35110	9331	1.1E+03
elongin B	1vcb	A	D	2.3	20571	37080	9574	9.5E+02
elongin C	1vcb	B	E	2.3	16285	29960	8495	1.0E+03
p53	1ycs	A	A	2.95	21487	37680	9599	7.7E+02
p53bp2	1ycs	B	B	2.95	21570	37090	9560	7.1E+02
CAMP dependent protein kinase	1ydr	A	E	2.35	38870	67210	15310	9.0E+02
protein kinase inhibitor peptide	1ydr	B	I	2.35	2219	5164	2362	1.1E+03
peroxisome proliferator receptor γ	2prg	A	A B	2.75	58933	103500	24380	1.3E+03
nuclear receptor coactivator src-1	2prg	B	C	2.75	3658	8739	3978	1.3E+03
transducin	2trc	A	B G	2	45331	77010	17850	2.1E+03
posducin	2trc	B	P	2	22803	44310	13940	2.2E+03
MHC and homologs complexed with others								
CD8	1akj	A	DE	2.4	25638	45510	11320	9.2E+02
MHC class I HLA-A2	1akj	B	ABC	2.4	44431	76950	18830	1.0E+03
T-cell receptor	1ao7	A	DE	2	35761	63660	16170	9.8E+02
MHC class I HLA-A2	1ao7	B	ABC	2	44445	77730	18980	9.9E+02
MHC class I H-2	1bqh	A	DEF	2.5	44168	76240	18660	1.4E+03
CD8	1bqh	B	KI	2.5	27709	50150	13300	1.4E+03
T-cell receptor D10	1d9k	A	EF	2.8	24468	42670	10810	1.0E+03
MHC class II I-Ak	1d9k	B	GHQ	2.8	44875	76430	18680	9.7E+02
KIR2dl2	1efx	A	DE	3.1	43262	76650	20260	7.3E+02
MHC class I HLA-CW3	1efx	B	ABC	3.1	44788	77580	19270	7.9E+02
MHC class I H-2Kb	1fo0	A	HLP	2.7	44437	76960	19170	6.5E+02
T-cell receptor BM3.3	1fo0	B	AB	2.7	25685	45000	11120	6.9E+02
MHC class II HLA-DR-1	1fyt	A	ABC	2.8	43149	74360	17580	1.0E+03
T-cell receptor	1fyt	B	DE	2.8	48985	85020	20450	1.0E+03
MHC class II HLA-DR-1	1hqr	A	ABC	1.9	40726	71100	17230	7.5E+02
Streptococcal pyrogenic exotoxin C	1hqr	B	D	1.9	23176	40650	10110	8.1E+02
FcRn	1i1a	A	AB	2.3	41398	72090	17510	7.1E+02
Fc	1i1a	B	CD	2.3	46193	81610	20720	7.1E+02
ly49a	1qo3	A	CD	3.5	29011	52230	14150	4.6E+02
MHC class I H-2Dd	1qo3	B	ABP	3.5	44604	77630	19210	5.2E+02
T-cell receptor	2ckb	A	AB	2.35	48453	84870	21330	9.5E+02
MHC class I H-2Kb	2ckb	B	HLP	2.35	44274	77160	19090	1.0E+03

	pdbid			Circularity	RMSD to plane	Atom burial	Atom accessibility
G-protein, cell cycle, signal transduction							
Giγ2	1got	A	A	1.5E-01	5.1E+00	9.2E+00	1.5E+01
Gtα-Gia	1got	B	B G	2.7E-01	4.4E+00	9.2E+00	1.4E+01
Gtβ	1gua	A	A	5.9E-01	1.4E+00	8.7E+00	1.4E+01
Gtγ	1gua	B	B	6.6E-01	1.2E+00	9.5E+00	1.7E+01
mcm1 transcriptiona regulator	1mnm	A	A B	3.5E-01	2.5E+00	6.7E+00	1.1E+01
mat-α-2 repressor	1mnm	B	C	3.5E-01	3.2E+00	1.0E+01	1.7E+01
ran	1rrp	A	A	3.3E-01	4.5E+00	9.1E+00	1.6E+01
nuclear core complex protein nup35	1rrp	B	B	3.0E-01	4.9E+00	9.4E+00	1.6E+01
p50-rhogap	1tx4	A	A	6.0E-01	2.8E+00	9.8E+00	1.6E+01
rhoa	1tx4	B	B	6.0E-01	2.7E+00	9.2E+00	1.5E+01
elongin B	1vcb	A	D	2.8E-01	2.3E+00	8.1E+00	1.3E+01
elongin C	1vcb	B	E	1.8E-01	2.2E+00	1.0E+01	1.6E+01
p53	1ycs	A	A	4.3E-01	2.9E+00	9.6E+00	1.7E+01
p53bp2	1ycs	B	B	3.1E-01	2.6E+00	7.6E+00	1.4E+01
CAMP dependent protein kinase	1ydr	A	E	5.1E-01	3.6E+00	7.5E+00	1.3E+01
protein kinase inhibitor peptide	1ydr	B	I	1.9E-01	2.4E+00	1.2E+01	1.8E+01
peroxisome proliferator receptor γ	2prg	A	A B	1.0E-01	4.6E+00	9.0E+00	1.4E+01
nuclear receptor coactivator src-1	2prg	B	C	8.4E-02	3.7E+00	1.1E+01	2.0E+01
transducin	2trc	A	B G	4.6E-01	4.1E+00	8.6E+00	1.3E+01
posducin	2trc	B	P	3.0E-01	4.1E+00	9.3E+00	1.5E+01
MHC and homologs complexed with others							
CD8	1akj	A	DE	2.9E-01	4.3E+00	8.7E+00	1.7E+01
MHC class I HLA-A2	1akj	B	ABC	2.5E-01	4.5E+00	9.2E+00	1.8E+01
T-cell receptor	1ao7	A	DE	3.9E-01	2.4E+00	8.5E+00	1.4E+01
MHC class I HLA-A2	1ao7	B	ABC	6.1E-01	2.7E+00	8.2E+00	1.3E+01
MHC class I H-2	1bqh	A	DEF	4.8E-01	4.1E+00	9.2E+00	1.5E+01
CD8	1bqh	B	KI	5.1E-01	4.0E+00	9.2E+00	1.5E+01
T-cell receptor D10	1d9k	A	EF	5.0E-01	2.0E+00	9.2E+00	1.5E+01
MHC class II I-Ak	1d9k	B	GHQ	9.4E-01	2.1E+00	8.5E+00	1.3E+01
KIR2dl2	1efx	A	DE	3.7E-01	1.5E+00	8.6E+00	1.3E+01
MHC class I HLA-CW3	1efx	B	ABC	4.8E-01	1.5E+00	1.1E+01	1.6E+01
MHC class I H-2Kb	1fo0	A	HLP	9.6E-01	2.1E+00	7.6E+00	1.4E+01
T-cell receptor BM3.3	1fo0	B	AB	5.0E-01	1.7E+00	9.6E+00	1.7E+01
MHC class II HLA-DR-1	1fyt	A	ABC	7.7E-01	2.4E+00	9.2E+00	1.4E+01
T-cell receptor	1fyt	B	DE	7.1E-01	2.2E+00	1.0E+01	1.5E+01
MHC class II HLA-DR-1	1hqr	A	ABC	6.9E-01	2.1E+00	8.0E+00	1.2E+01
Streptococcal pyrogenic exotoxin C	1hqr	B	D	3.9E-01	2.0E+00	9.0E+00	1.4E+01
FcRn	1i1a	A	AB	4.5E-01	2.6E+00	9.0E+00	1.4E+01
Fc	1i1a	B	CD	3.0E-01	2.5E+00	9.5E+00	1.4E+01
ly49a	1qo3	A	CD	3.0E-01	1.8E+00	9.5E+00	1.8E+01
MHC class I H-2Dd	1qo3	B	ABP	4.1E-01	1.8E+00	1.0E+01	1.9E+01
T-cell receptor	2ckb	A	AB	4.1E-01	1.7E+00	8.2E+00	1.4E+01
MHC class I H-2Kb	2ckb	B	HLP	6.0E-01	2.2E+00	8.7E+00	1.4E+01

	pdbid			Shape comple- mentarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
G-protein, cell cycle, signal transduction						
Giγ2	1got	A	A	7.7E-01	1.2E+00	1.6E+00
Gtα-Gia	1got	B	B G	7.7E-01	1.1E+00	1.6E+00
Gtβ	1gua	A	A	7.7E-01	1.0E+00	2.0E+00
Gtγ	1gua	B	B	7.7E-01	9.2E-01	2.0E+00
mcm1 transcriptiona regulator	1mnm	A	A B	6.6E-01	1.4E+00	2.0E+00
mat-α-2 repressor	1mnm	B	C	6.6E-01	1.1E+00	2.3E+00
ran	1rrp	A	A	7.0E-01	9.5E-01	2.0E+00
nuclear core complex protein nup35i	1rrp	B	B	7.0E-01	1.1E+00	2.1E+00
p50-rhogap	1tx4	A	A	7.3E-01	1.2E+00	1.7E+00
rhoa	1tx4	B	B	7.3E-01	1.0E+00	1.4E+00
elongin B	1vcb	A	D	6.6E-01	1.3E+00	1.9E+00
elongin C	1vcb	B	E	6.6E-01	1.3E+00	2.0E+00
p53	1ycs	A	A	7.6E-01	1.0E+00	2.1E+00
p53bp2	1ycs	B	B	7.6E-01	1.1E+00	1.8E+00
CAMP dependent protein kinase	1ydr	A	E	7.2E-01	1.2E+00	1.8E+00
protein kinase inhibitor peptide	1ydr	B	I	7.2E-01	1.4E+00	2.6E+00
peroxisome proliferator receptor γ	2prg	A	A B	7.3E-01	1.3E+00	1.9E+00
nuclear receptor coactivator src-1	2prg	B	C	7.3E-01	1.5E+00	2.7E+00
transducin	2trc	A	B G	7.1E-01	1.1E+00	1.7E+00
posducin	2trc	B	P	7.1E-01	1.1E+00	1.9E+00
MHC and homologs complexed with others						
CD8	1akj	A	DE	6.6E-01	1.2E+00	2.2E+00
MHC class I HLA-A2	1akj	B	ABC	6.6E-01	9.1E-01	2.2E+00
T-cell receptor	1ao7	A	DE	6.3E-01	1.0E+00	2.0E+00
MHC class I HLA-A2	1ao7	B	ABC	6.3E-01	1.1E+00	1.9E+00
MHC class I H-2	1bqh	A	DEF	6.0E-01	1.0E+00	2.2E+00
CD8	1bqh	B	KI	6.0E-01	1.0E+00	1.9E+00
T-cell receptor D10	1d9k	A	EF	6.1E-01	9.9E-01	2.0E+00
MHC class II I-Ak	1d9k	B	GHQ	6.1E-01	1.1E+00	2.0E+00
KIR2dl2	1efx	A	DE	6.8E-01	1.3E+00	1.8E+00
MHC class I HLA-CW3	1efx	B	ABC	6.8E-01	1.0E+00	1.9E+00
MHC class I H-2Kb	1fo0	A	HLP	6.0E-01	1.3E+00	1.7E+00
T-cell receptor BM3.3	1fo0	B	AB	6.0E-01	1.4E+00	1.7E+00
MHC class II HLA-DR-1	1fyt	A	ABC	5.6E-01	9.9E-01	2.0E+00
T-cell receptor	1fyt	B	DE	5.6E-01	9.0E-01	1.6E+00
MHC class II HLA-DR-1	1hqr	A	ABC	6.9E-01	1.1E+00	2.1E+00
Streptococcal pyrogenic exotoxin C	1hqr	B	D	6.9E-01	1.1E+00	1.9E+00
FcRn	1i1a	A	AB	7.2E-01	1.1E+00	2.1E+00
Fc	1i1a	B	CD	7.2E-01	1.0E+00	2.3E+00
ly49a	1qo3	A	CD	7.7E-01	1.0E+00	1.7E+00
MHC class I H-2Dd	1qo3	B	ABP	7.7E-01	9.1E-01	1.6E+00
T-cell receptor	2ckb	A	AB	4.1E-01	9.3E-01	1.9E+00
MHC class I H-2Kb	2ckb	B	HLP	4.1E-01	1.2E+00	1.9E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
G-protein, cell cycle, signal transduction							
Giγ2	1got	A	A	8.0E+00	6.4E-01	2.6E+01	2.1E+00
Gtα-Gia	1got	B	B G	8.0E+00	6.6E-01	2.6E+01	2.1E+00
Gtβ	1gua	A	A	1.3E+01	2.0E+00	2.1E+01	3.3E+00
Gtγ	1gua	B	B	1.3E+01	2.1E+00	2.1E+01	3.4E+00
mcm1 transcriptiona regulator	1mnm	A	A B	1.3E+01	2.1E+00	2.1E+01	3.4E+00
mat-α-2 repressor	1mnm	B	C	1.3E+01	1.1E+00	2.1E+01	1.8E+00
ran	1rrp	A	A	1.0E+00	4.2E-02	4.0E+00	1.7E-01
nuclear core complex protein nup35	1rrp	B	B	1.0E+00	4.3E-02	4.0E+00	1.7E-01
p50-rhogap	1tx4	A	A	8.0E+00	7.2E-01	1.3E+01	1.2E+00
rhoa	1tx4	B	B	8.0E+00	7.0E-01	1.3E+01	1.1E+00
elongin B	1vcb	A	D	1.1E+01	1.2E+00	1.7E+01	1.8E+00
elongin C	1vcb	B	E	1.1E+01	1.1E+00	1.7E+01	1.6E+00
p53	1ycs	A	A	1.2E+01	1.6E+00	2.0E+01	2.6E+00
p53bp2	1ycs	B	B	1.2E+01	1.7E+00	2.0E+01	2.8E+00
CAMP dependent protein kinase	1ydr	A	E	3.7E+01	4.1E+00	6.3E+01	7.0E+00
protein kinase inhibitor peptide	1ydr	B	I	3.7E+01	3.5E+00	6.3E+01	5.9E+00
peroxisome proliferator receptor γ	2prg	A	A B	1.0E+01	7.9E-01	2.2E+01	1.7E+00
nuclear receptor coactivator src-1	2prg	B	C	1.0E+01	7.5E-01	2.2E+01	1.7E+00
transducin	2trc	A	B G	2.0E+01	9.4E-01	4.5E+01	2.1E+00
posducin	2trc	B	P	2.0E+01	8.9E-01	4.5E+01	2.0E+00
MHC and homologs complexed with others							
CD8	1akj	A	DE	1.3E+01	1.4E+00	2.3E+01	2.5E+00
MHC class I HLA-A2	1akj	B	ABC	1.3E+01	1.2E+00	2.3E+01	2.2E+00
T-cell receptor	1ao7	A	DE	1.4E+01	1.4E+00	2.1E+01	2.1E+00
MHC class I HLA-A2	1ao7	B	ABC	1.4E+01	1.4E+00	2.1E+01	2.1E+00
MHC class I H-2	1bqh	A	DEF	8.0E+00	5.5E-01	1.3E+01	9.0E-01
CD8	1bqh	B	KI	8.0E+00	5.6E-01	1.3E+01	9.0E-01
T-cell receptor D10	1d9k	A	EF	6.0E+00	5.9E-01	1.8E+01	1.8E+00
MHC class II I-Ak	1d9k	B	GHQ	6.0E+00	6.2E-01	1.8E+01	1.8E+00
KIR2dl2	1efx	A	DE	1.9E+01	2.6E+00	3.5E+01	4.8E+00
MHC class I HLA-CW3	1efx	B	ABC	1.9E+01	2.4E+00	3.5E+01	4.4E+00
MHC class I H-2Kb	1fo0	A	HLP	1.4E+01	2.1E+00	2.0E+01	3.1E+00
T-cell receptor BM3.3	1fo0	B	AB	1.4E+01	2.0E+00	2.0E+01	2.9E+00
MHC class II HLA-DR-1	1fyt	A	ABC	8.0E+00	7.9E-01	1.5E+01	1.5E+00
T-cell receptor	1fyt	B	DE	8.0E+00	7.7E-01	1.5E+01	1.4E+00
MHC class II HLA-DR-1	1hqr	A	ABC	1.3E+01	1.7E+00	2.3E+01	3.1E+00
Streptococcal pyrogenic exotoxin C	1hqr	B	D	1.3E+01	1.6E+00	2.3E+01	2.8E+00
FcRn	1i1a	A	AB	2.1E+01	3.0E+00	3.5E+01	4.9E+00
Fc	1i1a	B	CD	2.1E+01	3.0E+00	3.5E+01	4.9E+00
ly49a	1qo3	A	CD	1.9E+01	4.1E+00	5.1E+01	1.1E+01
MHC class I H-2Dd	1qo3	B	ABP	1.9E+01	3.7E+00	5.1E+01	9.8E+00
T-cell receptor	2ckb	A	AB	7.5E+01	7.9E+00	1.3E+02	1.3E+01
MHC class I H-2Kb	2ckb	B	HLP	7.5E+01	7.5E+00	1.3E+02	1.2E+01

	pdbid			%non-pola	%polar	%charged
G-protein, cell cycle, signal transduction						
Giγ2	1got	A	A	6.2E-01	2.0E-01	1.8E-01
Gtα-Gia	1got	B	B G	5.6E-01	2.9E-01	1.5E-01
Gtβ	1gua	A	A	4.8E-01	2.1E-01	3.1E-01
Gtγ	1gua	B	B	4.7E-01	3.4E-01	1.9E-01
mcm1 transcriptiona regulator	1mnm	A	A B	6.9E-01	2.0E-01	1.2E-01
mat-α-2 repressor	1mnm	B	C	6.1E-01	2.8E-01	1.1E-01
ran	1rrp	A	A	6.5E-01	2.4E-01	1.2E-01
nuclear core complex protein nup35	1rrp	B	B	6.4E-01	1.7E-01	1.9E-01
p50-rhogap	1tx4	A	A	5.8E-01	2.5E-01	1.7E-01
rhoa	1tx4	B	B	6.2E-01	2.5E-01	1.3E-01
elongin B	1vcb	A	D	6.7E-01	2.2E-01	1.1E-01
elongin C	1vcb	B	E	7.8E-01	1.7E-01	5.6E-02
p53	1ycs	A	A	4.2E-01	3.0E-01	2.8E-01
p53bp2	1ycs	B	B	5.0E-01	3.4E-01	1.6E-01
CAMP dependent protein kinase	1ydr	A	E	5.7E-01	1.7E-01	2.6E-01
protein kinase inhibitor peptide	1ydr	B	I	6.1E-01	1.5E-01	2.4E-01
peroxisome proliferator receptor γ	2prg	A	A B	7.0E-01	5.9E-02	2.4E-01
nuclear receptor coactivator src-1	2prg	B	C	7.9E-01	1.7E-01	4.2E-02
transducin	2trc	A	B G	4.7E-01	2.9E-01	2.5E-01
posducin	2trc	B	P	6.2E-01	2.0E-01	1.9E-01
MHC and homologs complexed with others						
CD8	1akj	A	DE	3.8E-01	4.4E-01	1.7E-01
MHC class I HLA-A2	1akj	B	ABC	4.4E-01	2.9E-01	2.7E-01
T-cell receptor	1ao7	A	DE	5.8E-01	2.7E-01	1.4E-01
MHC class I HLA-A2	1ao7	B	ABC	5.8E-01	2.4E-01	1.8E-01
MHC class I H-2	1bqh	A	DEF	5.1E-01	2.5E-01	2.4E-01
CD8	1bqh	B	KI	4.4E-01	4.3E-01	1.3E-01
T-cell receptor D10	1d9k	A	EF	5.5E-01	2.7E-01	1.7E-01
MHC class II I-Ak	1d9k	B	GHQ	5.3E-01	2.5E-01	2.3E-01
KIR2dl2	1efx	A	DE	5.7E-01	2.5E-01	1.8E-01
MHC class I HLA-CW3	1efx	B	ABC	5.1E-01	1.3E-01	3.6E-01
MHC class I H-2Kb	1fo0	A	HLP	5.4E-01	2.4E-01	2.2E-01
T-cell receptor BM3.3	1fo0	B	AB	4.5E-01	2.9E-01	2.6E-01
MHC class II HLA-DR-1	1fyt	A	ABC	6.2E-01	2.0E-01	1.8E-01
T-cell receptor	1fyt	B	DE	6.3E-01	1.8E-01	1.9E-01
MHC class II HLA-DR-1	1hqr	A	ABC	6.8E-01	1.8E-01	1.4E-01
Streptococcal pyrogenic exotoxin C	1hqr	B	D	4.6E-01	4.6E-01	7.4E-02
FcRn	1i1a	A	AB	6.1E-01	1.6E-01	2.3E-01
Fc	1i1a	B	CD	5.8E-01	3.3E-01	9.6E-02
ly49a	1qo3	A	CD	3.4E-01	2.5E-01	4.1E-01
MHC class I H-2Dd	1qo3	B	ABP	3.7E-01	1.5E-01	4.8E-01
T-cell receptor	2ckb	A	AB	6.0E-01	3.7E-01	3.1E-02
MHC class I H-2Kb	2ckb	B	HLP	5.6E-01	1.9E-01	2.5E-01

	pdbid			Reso-lution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Fc complexes								
Fc IgG4	1adq	A	A	2.36	23413	43330	12220	6.7E+02
rheumatoid factor	1adq	B	H L	2.36	47109	82750	20360	7.5E+02
Fcγ RIII	1e4k	A	C	2.8	48738	87270	22830	8.0E+02
Fc IgG1	1e4k	B	A B	2.8	19541	35090	9582	8.1E+02
FcεRI	1f6a	A	A	1.8	19826	36750	10330	8.7E+02
Fc IgE	1f6a	B	B D	1.8	48441	89290	24110	8.2E+02
fragment B of protein A	1fc2	A	C	1.9	4990	9848	3294	6.3E+02
Fc IgG	1fc2	B	D	1.9	23445	43090	12120	6.6E+02
Fc IgG	1fcc	A	A	1.8	23446	42950	12010	6.6E+02
protein G	1fcc	B	C	1.8	6133	11770	3738	6.8E+02
Cell surface receptor ligand								
fibroblast GF2	1cvs	A	A B	2.6	14715	25880	6530	1.7E+03
fibroblast-GFR1	1cvs	B	C	2.6	43512	79200	22590	1.7E+03
death receptor-5	1d0g	A	A B D	2	52731	90200	20280	1.5E+03
apoptosis-2 ligand	1d0g	B	R	2	12058	22660	7559	1.5E+03
fibroblast GF1	1e0o	A	A	2.5	13847	24590	6302	8.6E+02
fibroblast GFR2	1e0o	B	B	2.5	21302	61790	11440	8.9E+02
erythropoietin receptor	1ebp	A	A B	1.8	46500	86040	24220	9.4E+02
epo mimetic peptide	1ebp	B	C D	1.8	3694	7399	2654	9.7E+02
erythropoietin	1eer	A	A	2.3	18330	33850	9564	1.6E+03
erythropoietin receptor	1eer	B	B C	2.3	47092	84110	22630	1.7E+03
bone-morphogenetic protein-2	1es7	A	A C	1.8	23180	41620	11170	1.1E+03
bone-morphogenetic protein recepto	1es7	B	B	1.8	9246	17780	5922	1.1E+03
fibroblast-GF2	1ev2	A	A	2.85	13730	24980	6443	1.3E+03
fibroblast-GFR2	1ev2	B	E	2.85	20403	38660	11390	1.3E+03
prolactin receptor	1f6f	A	B C	2.2	42805	78240	20810	1.8E+03
placental lactogen	1f6f	B	A	2.2	20925	38940	10700	1.7E+03
interferon γ	1fg9	A	A B	1.8	29592	52370	13510	9.4E+02
interferon γ receptor α chain	1fg9	B	D	1.8	22991	41630	11640	8.9E+02
tnf receptor associated factor 3	1fll	A	A	2.7	23107	43280	12990	5.6E+02
CD40	1fll	B	X	2.7	2170	5313	2576	6.2E+02
fibroblast-GF2	1fq9	A	A	2	29430	25840	6567	1.7E+03
fibroblast-GFR1	1fq9	B	C D	2	23207	81580	23250	1.7E+03
gp120	1gc1	A	H L G	2.8	80503	138400	22040	9.7E+02
CD4	1gc1	B	C	2.8	20049	35730	32150	9.4E+02
growth hormone	1hwg	A	A	2.6	20626	37090	9670	2.1E+03
growth hormone binding protein	1hwg	B	B C	2.6	43682	76960	19320	2.1E+03
interleukin-4	1iar	A	A	3.3	14939	27300	7408	7.4E+02
interleukin-4 receptor α	1iar	B	B	3.3	21457	38810	10780	8.0E+02
T-cell receptor 14.3.D	1jck	A	A C	2.4	51584	89870	22920	2.1E+03
Staphylococcal enterotoxin C3	1jck	B	B D	2.4	55124	94700	22780	2.1E+03
CD2	1qa9	A	A	2.8	11962	23000	6670	7.0E+02
CD58	1qa9	B	B	2.8	10965	20560	5990	6.3E+02
vascular endothelial growth factor	1qty	A	V W	2.2	22148	39390	10880	7.8E+02
FMS-like tyrosine kinase-1	1qty	B	X	2.2	10756	19990	5747	8.4E+02
tumor necrosis factor receptor p55	1tnr	A	A	2.6	15834	28940	7768	5.7E+02
tumor necrosis factor beta	1tnr	B	R	2.6	15661	29570	9568	6.3E+02
nerve growth factor	1www	A	V W	2.6	24485	45090	12920	2.1E+03
TRKa receptor	1www	B	X Y	2.6	22130	40650	11940	2.2E+03

	pdbid			Circularity	RMSD to plane	Atom burial	Atom accessibility
Fc complexes							
Fc IgG4	1adq	A	A	3.6E-01	2.0E+00	8.4E+00	1.5E+01
rheumatoid factor	1adq	B	H L	5.5E-01	1.9E+00	9.3E+00	1.6E+01
FcγRIII	1e4k	A	C	2.4E-01	3.3E+00	8.8E+00	1.7E+01
Fc IgG1	1e4k	B	A B	3.0E-01	3.4E+00	8.4E+00	1.5E+01
FcεRI	1f6a	A	A	3.2E-01	3.6E+00	8.9E+00	1.4E+01
Fc IgE	1f6a	B	B D	2.3E-01	4.0E+00	7.3E+00	1.4E+01
fragment B of protein A	1fc2	A	C	5.6E-01	2.1E+00	8.8E+00	1.4E+01
Fc IgG	1fc2	B	D	5.3E-01	2.5E+00	8.9E+00	1.3E+01
Fc IgG	1fcc	A	A	6.7E-01	2.0E+00	9.4E+00	1.5E+01
protein G	1fcc	B	C	8.5E-01	1.9E+00	9.5E+00	1.4E+01
Cell surface receptor ligand							
fibroblast GF2	1cvs	A	A B	4.7E-01	4.9E+00	9.5E+00	1.4E+01
fibroblast-GFR1	1cvs	B	C	3.4E-01	5.1E+00	8.8E+00	1.4E+01
death receptor-5	1d0g	A	A B D	2.8E-01	3.3E+00	8.7E+00	1.6E+01
apoptosis-2 ligand	1d0g	B	R	1.7E-01	3.4E+00	1.1E+01	1.7E+01
fibroblast GF1	1e0o	A	A	2.8E-01	4.0E+00	8.9E+00	1.5E+01
fibroblast GFR2	1e0o	B	B	1.5E-01	3.9E+00	9.7E+00	1.7E+01
erythropoietin receptor	1ebp	A	A B	3.8E-01	5.1E+00	9.5E+00	1.5E+01
epo mimetic peptide	1ebp	B	C D	5.3E-01	4.5E+00	9.6E+00	1.5E+01
erythropoietin	1eer	A	A	4.3E-01	5.0E+00	9.5E+00	1.5E+01
erythropoietin receptor	1eer	B	B C	2.8E-01	4.8E+00	9.3E+00	1.4E+01
bone-morphogenetic protein-2	1es7	A	A C	5.2E-01	3.5E+00	9.3E+00	1.4E+01
bone-morphogenetic protein recepto	1es7	B	B	4.7E-01	2.7E+00	9.4E+00	1.5E+01
fibroblast-GF2	1ev2	A	A	3.1E-01	4.4E+00	9.7E+00	1.5E+01
fibroblast-GFR2	1ev2	B	E	2.8E-01	4.0E+00	9.3E+00	1.5E+01
prolactin receptor	1f6f	A	B C	1.2E-01	4.8E+00	9.1E+00	1.4E+01
placental lactogen	1f6f	B	A	1.8E-01	4.3E+00	9.2E+00	1.4E+01
interferon γ	1fg9	A	A B	3.4E-01	2.1E+00	9.6E+00	1.6E+01
interferon γ receptor α chain	1fg9	B	D	2.0E-01	2.0E+00	7.5E+00	1.3E+01
tnf receptor associated factor 3	1fll	A	A	2.2E-01	2.7E+00	7.2E+00	1.4E+01
CD40	1fll	B	X	1.0E-01	1.9E+00	1.0E+01	1.9E+01
fibroblast-GF2	1fq9	A	A	4.6E-01	4.9E+00	9.6E+00	1.4E+01
fibroblast-GFR1	1fq9	B	C D	2.9E-01	5.0E+00	8.9E+00	1.4E+01
gp120	1gc1	A	H L G	2.3E-01	4.7E+00	8.3E+00	1.8E+01
CD4	1gc1	B	C	2.8E-01	3.9E+00	9.8E+00	1.5E+04
growth hormone	1hwg	A	A	2.4E-01	5.5E+00	8.9E+00	1.4E+01
growth hormone binding protein	1hwg	B	B C	1.6E-01	5.9E+00	8.7E+00	1.4E+01
interleukin-4	1iar	A	A	5.6E-01	2.3E+00	9.2E+00	1.3E+01
interleukin-4 receptor α	1iar	B	B	3.8E-01	2.4E+00	9.9E+00	1.5E+01
T-cell receptor 14.3.D	1jck	A	A C	1.1E-01	7.4E+00	9.7E+00	1.7E+01
Staphylococcal enterotoxin C3	1jck	B	B D	1.0E-01	7.1E+00	8.1E+00	1.6E+01
CD2	1qa9	A	A	5.1E-01	1.9E+00	1.0E+01	1.8E+01
CD58	1qa9	B	B	6.6E-01	2.0E+00	9.1E+00	1.7E+01
vascular endothelial growth factor	1qty	A	V W	6.0E-01	2.0E+00	8.7E+00	1.6E+01
FMS-like tyrosine kinase-1	1qty	B	X	6.1E-01	1.8E+00	1.1E+01	1.7E+01
tumor necrosis factor receptor p55	1tnr	A	A	1.2E-01	2.4E+00	8.2E+00	1.6E+01
tumor necrosis factor beta	1tnr	B	R	1.0E-01	2.3E+00	9.4E+00	1.7E+01
nerve growth factor	1www	A	V W	5.7E-01	5.6E+00	8.2E+00	1.4E+01
TRKa receptor	1www	B	X Y	2.6E-01	5.5E+00	1.0E+01	1.6E+01

	pdbid			Shape complementarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Fc complexes						
Fc IgG4	1adq	A	A	6.9E-01	1.1E+00	2.2E+00
rheumatoid factor	1adq	B	H L	6.9E-01	8.8E-01	2.1E+00
FcγRIII	1e4k	A	C	6.7E-01	9.7E-01	2.2E+00
Fc IgG1	1e4k	B	A B	6.7E-01	9.6E-01	2.2E+00
FcεRI	1f6a	A	A	6.9E-01	1.0E+00	2.0E+00
Fc IgE	1f6a	B	B D	6.9E-01	1.0E+00	1.9E+00
fragment B of protein A	1fc2	A	C	6.6E-01	1.2E+00	1.9E+00
Fc IgG	1fc2	B	D	6.6E-01	1.4E+00	1.8E+00
Fc IgG	1fcc	A	A	5.7E-01	9.8E-01	1.7E+00
protein G	1fcc	B	C	5.7E-01	1.1E+00	1.7E+00
Cell surface receptor ligand						
fibroblast GF2	1cvs	A	A B	6.5E-01	1.1E+00	2.1E+00
fibroblast-GFR1	1cvs	B	C	6.5E-01	1.0E+00	2.1E+00
death receptor-5	1d0g	A	A B D	6.9E-01	1.1E+00	1.7E+00
apoptosis-2 ligand	1d0g	B	R	6.9E-01	1.4E+00	2.2E+00
fibroblast GF1	1e0o	A	A	6.6E-01	1.1E+00	2.0E+00
fibroblast GFR2	1e0o	B	B	6.6E-01	1.1E+00	2.3E+00
erythropoietin receptor	1ebp	A	A B	7.5E-01	1.3E+00	2.5E+00
epo mimetic peptide	1ebp	B	C D	7.5E-01	1.0E+00	2.0E+00
erythropoietin	1eer	A	A	6.5E-01	1.1E+00	1.8E+00
erythropoietin receptor	1eer	B	B C	6.5E-01	1.2E+00	1.6E+00
bone-morphogenetic protein-2	1es7	A	A C	6.9E-01	1.3E+00	2.0E+00
bone-morphogenetic protein recepto	1es7	B	B	6.9E-01	1.1E+00	2.0E+00
fibroblast-GF2	1ev2	A	A	7.0E-01	1.1E+00	1.9E+00
fibroblast-GFR2	1ev2	B	E	7.0E-01	1.1E+00	2.0E+00
prolactin receptor	1f6f	A	B C	6.9E-01	1.0E+00	1.9E+00
placental lactogen	1f6f	B	A	6.9E-01	1.2E+00	2.0E+00
interferon γ	1fg9	A	A B	7.2E-01	1.1E+00	2.1E+00
interferon γ receptor α chain	1fg9	B	D	7.2E-01	1.0E+00	2.1E+00
tnf receptor associated factor 3	1fll	A	A	6.0E-01	1.2E+00	1.7E+00
CD40	1fll	B	X	6.0E-01	1.2E+00	1.8E+00
fibroblast-GF2	1fq9	A	A	6.8E-01	1.0E+00	1.3E+00
fibroblast-GFR1	1fq9	B	C D	6.8E-01	1.0E+00	1.4E+00
gp120	1gc1	A	H L G	7.2E-01	1.1E+00	2.1E+00
CD4	1gc1	B	C	7.2E-01	1.1E+00	2.1E+00
growth hormone	1hwg	A	A	6.9E-01	1.2E+00	1.9E+00
growth hormone binding protein	1hwg	B	B C	6.9E-01	1.1E+00	1.9E+00
interleukin-4	1iar	A	A	7.4E-01	1.0E+00	1.8E+00
interleukin-4 receptor α	1iar	B	B	7.4E-01	1.1E+00	1.5E+00
T-cell receptor 14.3.D	1jck	A	A C	5.2E-01	1.0E+00	2.4E+00
Staphylococcal enterotoxin C3	1jck	B	B D	5.2E-01	9.0E-01	2.3E+00
CD2	1qa9	A	A	5.6E-01	1.1E+00	2.3E+00
CD58	1qa9	B	B	5.6E-01	1.1E+00	2.1E+00
vascular endothelial growth factor	1qty	A	V W	6.1E-01	1.3E+00	2.2E+00
FMS-like tyrosine kinase-1	1qty	B	X	6.1E-01	1.5E+00	2.1E+00
tumor necrosis factor receptor p55	1tnr	A	A	6.3E-01	1.0E+00	2.1E+00
tumor necrosis factor beta	1tnr	B	R	6.3E-01	9.6E-01	2.5E+00
nerve growth factor	1www	A	V W	7.1E-01	1.2E+00	1.8E+00
TRKa receptor	1www	B	X Y	7.1E-01	1.1E+00	1.8E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
Fc complexes							
Fc IgG4	1adq	A	A	7.0E+00	1.0E+00	1.0E+01	1.5E+00
rheumatoid factor	1adq	B	H L	7.0E+00	9.3E-01	1.0E+01	1.3E+00
FcγRIII	1e4k	A	C	8.0E+00	9.9E-01	2.3E+01	2.9E+00
Fc IgG1	1e4k	B	A B	8.0E+00	9.9E-01	2.3E+01	2.8E+00
FcεRI	1f6a	A	A	1.0E+00	1.2E-01	1.0E+00	1.2E-01
Fc IgE	1f6a	B	B D	1.0E+00	1.2E-01	1.0E+00	1.2E-01
fragment B of protein A	1fc2	A	C	1.4E+01	2.2E+00	2.7E+01	4.3E+00
Fc IgG	1fc2	B	D	1.4E+01	2.1E+00	2.7E+01	4.1E+00
Fc IgG	1fcc	A	A	1.0E+01	1.5E+00	2.0E+01	3.0E+00
protein G	1fcc	B	C	1.0E+01	1.5E+00	2.0E+01	2.9E+00
Cell surface receptor ligand							
fibroblast GF2	1cvs	A	A B	5.0E+00	3.0E-01	7.0E+00	4.2E-01
fibroblast-GFR1	1cvs	B	C	5.0E+00	3.0E-01	7.0E+00	4.2E-01
death receptor-5	1d0g	A	A B D	1.7E+01	1.1E+00	2.7E+01	1.8E+00
apoptosis-2 ligand	1d0g	B	R	1.7E+01	1.1E+00	2.7E+01	1.8E+00
fibroblast GF1	1e0o	A	A	1.2E+01	1.4E+00	2.3E+01	2.7E+00
fibroblast GFR2	1e0o	B	B	1.2E+01	1.3E+00	2.3E+01	2.6E+00
erythropoietin receptor	1ebp	A	A B	7.0E+00	7.5E-01	1.4E+01	1.5E+00
epo mimetic peptide	1ebp	B	C D	7.0E+00	7.2E-01	1.4E+01	1.4E+00
erythropoietin	1eer	A	A	1.1E+01	6.7E-01	1.9E+01	1.2E+00
erythropoietin receptor	1eer	B	B C	1.1E+01	6.6E-01	1.9E+01	1.1E+00
bone-morphogenetic protein-2	1es7	A	A C	1.4E+01	1.3E+00	2.4E+01	2.2E+00
bone-morphogenetic protein recepto	1es7	B	B	1.4E+01	1.2E+00	2.4E+01	2.1E+00
fibroblast-GF2	1ev2	A	A	6.0E+00	4.5E-01	1.0E+01	7.5E-01
fibroblast-GFR2	1ev2	B	E	6.0E+00	4.5E-01	1.0E+01	7.4E-01
prolactin receptor	1f6f	A	B C	1.0E+01	5.7E-01	2.1E+01	1.2E+00
placental lactogen	1f6f	B	A	1.0E+01	5.8E-01	2.1E+01	1.2E+00
interferon γ	1fg9	A	A B	8.0E+00	8.5E-01	1.3E+01	1.4E+00
interferon γ receptor α chain	1fg9	B	D	8.0E+00	8.9E-01	1.3E+01	1.5E+00
tnf receptor associated factor 3	1fll	A	A	7.0E+00	1.2E+00	1.4E+01	2.5E+00
CD40	1fll	B	X	7.0E+00	1.1E+00	1.4E+01	2.3E+00
fibroblast-GF2	1fq9	A	A	1.8E+01	1.1E+00	3.5E+01	2.1E+00
fibroblast-GFR1	1fq9	B	C D	1.8E+01	1.1E+00	3.5E+01	2.1E+00
gp120	1gc1	A	H L G	1.5E+01	1.5E+00	2.6E+01	2.7E+00
CD4	1gc1	B	C	1.5E+01	1.6E+00	2.6E+01	2.8E+00
growth hormone	1hwg	A	A	8.0E+00	3.9E-01	1.0E+01	4.8E-01
growth hormone binding protein	1hwg	B	B C	8.0E+00	3.8E-01	1.0E+01	4.8E-01
interleukin-4	1iar	A	A	9.0E+00	1.2E+00	1.6E+01	2.2E+00
interleukin-4 receptor α	1iar	B	B	9.0E+00	1.1E+00	1.6E+01	2.0E+00
T-cell receptor 14.3.D	1jck	A	A C	9.0E+00	4.2E-01	1.6E+01	7.5E-01
Staphylococcal enterotoxin C3	1jck	B	B D	9.0E+00	4.4E-01	1.6E+01	7.8E-01
CD2	1qa9	A	A	7.0E+00	1.0E+00	1.4E+01	2.0E+00
CD58	1qa9	B	B	7.0E+00	1.1E+00	1.4E+01	2.2E+00
vascular endothelial growth factor	1qty	A	V W	1.6E+01	2.1E+00	3.7E+01	4.8E+00
FMS-like tyrosine kinase-1	1qty	B	X	1.6E+01	1.9E+00	3.7E+01	4.4E+00
tumor necrosis factor receptor p55	1tnr	A	A	1.6E+01	2.8E+00	2.3E+01	4.0E+00
tumor necrosis factor beta	1tnr	B	R	1.6E+01	2.5E+00	2.3E+01	3.7E+00
nerve growth factor	1www	A	V W	6.0E+00	2.9E-01	1.3E+01	6.3E-01
TRKa receptor	1www	B	X Y	6.0E+00	2.8E-01	1.3E+01	6.0E-01

	pdbid			%non-pola	%polar	%charged
Fc complexes						
Fc IgG4	1adq	A	A	5.2E-01	4.4E-01	4.0E-02
rheumatoid factor	1adq	B	H L	6.0E-01	2.4E-01	1.6E-01
FcγRIII	1e4k	A	C	6.6E-01	2.9E-01	5.4E-02
Fc IgG1	1e4k	B	A B	6.8E-01	1.9E-01	1.2E-01
FcεRI	1f6a	A	A	6.3E-01	2.7E-01	9.6E-02
Fc IgE	1f6a	B	B D	6.1E-01	2.5E-01	1.4E-01
fragment B of protein A	1fc2	A	C	6.2E-01	3.5E-01	3.4E-02
Fc IgG	1fc2	B	D	6.0E-01	3.7E-01	2.8E-02
Fc IgG	1fcc	A	A	5.6E-01	3.7E-01	7.3E-02
protein G	1fcc	B	C	5.2E-01	2.3E-01	2.5E-01
Cell surface receptor ligand						
fibroblast GF2	1cvs	A	A B	6.0E-01	2.7E-01	1.4E-01
fibroblast-GFR1	1cvs	B	C	6.8E-01	1.9E-01	1.3E-01
death receptor-5	1d0g	A	A B D	4.6E-01	3.4E-01	2.0E-01
apoptosis-2 ligand	1d0g	B	R	5.6E-01	1.4E-01	2.9E-01
fibroblast GF1	1e0o	A	A	5.6E-01	2.4E-01	2.0E-01
fibroblast GFR2	1e0o	B	B	6.3E-01	1.3E-01	2.4E-01
erythropoietin receptor	1ebp	A	A B	7.0E-01	3.0E-01	9.1E-04
epo mimetic peptide	1ebp	B	C D	7.1E-01	2.9E-01	0.0E+00
erythropoietin	1eer	A	A	5.4E-01	1.6E-01	3.0E-01
erythropoietin receptor	1eer	B	B C	5.7E-01	2.7E-01	1.6E-01
bone-morphogenetic protein-2	1es7	A	A C	6.6E-01	2.8E-01	6.6E-02
bone-morphogenetic protein recepto	1es7	B	B	6.0E-01	1.9E-01	2.1E-01
fibroblast-GF2	1ev2	A	A	6.1E-01	2.4E-01	1.4E-01
fibroblast-GFR2	1ev2	B	E	6.3E-01	2.3E-01	1.4E-01
prolactin receptor	1f6f	A	B C	6.4E-01	1.9E-01	1.7E-01
placental lactogen	1f6f	B	A	5.7E-01	2.4E-01	1.8E-01
interferon γ	1fg9	A	A B	5.6E-01	2.9E-01	1.5E-01
interferon γ receptor α chain	1fg9	B	D	5.3E-01	3.4E-01	1.3E-01
tnf receptor associated factor 3	1fll	A	A	5.7E-01	2.3E-01	2.0E-01
CD40	1fll	B	X	6.3E-01	2.6E-01	1.1E-01
fibroblast-GF2	1fq9	A	A	5.9E-01	2.6E-01	1.5E-01
fibroblast-GFR1	1fq9	B	C D	6.9E-01	1.9E-01	1.2E-01
gp120	1gc1	A	H L G	6.1E-01	2.7E-01	1.2E-01
CD4	1gc1	B	C	6.2E-01	2.8E-01	9.9E-02
growth hormone	1hwg	A	A	6.1E-01	1.9E-01	2.0E-01
growth hormone binding protein	1hwg	B	B C	5.8E-01	2.5E-01	1.7E-01
interleukin-4	1iar	A	A	4.9E-01	1.6E-01	3.6E-01
interleukin-4 receptor α	1iar	B	B	6.6E-01	1.9E-01	1.4E-01
T-cell receptor 14.3.D	1jck	A	A C	5.1E-01	4.0E-01	9.2E-02
Staphylococcal enterotoxin C3	1jck	B	B D	6.0E-01	3.2E-01	7.9E-02
CD2	1qa9	A	A	4.1E-01	1.9E-01	4.1E-01
CD58	1qa9	B	B	4.7E-01	7.7E-02	4.6E-01
vascular endothelial growth factor	1qty	A	V W	6.6E-01	2.4E-01	9.7E-02
FMS-like tyrosine kinase-1	1qty	B	X	7.5E-01	1.6E-01	9.2E-02
tumor necrosis factor receptor p55	1tnr	A	A	5.5E-01	2.9E-01	1.5E-01
tumor necrosis factor beta	1tnr	B	R	5.5E-01	3.0E-01	1.5E-01
nerve growth factor	1www	A	V W	6.2E-01	2.3E-01	1.5E-01
TRKa receptor	1www	B	X Y	6.7E-01	2.7E-01	6.8E-02

	pdbid			Reso-lution A	Protein mass Daltons	Average volume A^3	Total surface area A^2	interface area A^2
Miscellaneous								
cyclophilin A	1ak4	A	A	1.75	17977	30970	7470	4.6E+02
HIV-1 capsid	1ak4	B	D	1.75	16073	29560	8551	5.6E+02
nucleotide exchangefactor GRPE	1dkg	A	A B	2.5	33470	62110	17770	9.7E+02
molecular chaperone DNAK	1dkg	B	D	2.5	39886	69660	16680	9.8E+02
Fab IgG1	1igc	A	L H	2.1	47259	80670	18760	6.3E+02
protein G	1igc	B	A	2.1	6303	12000	3781	6.8E+02
transthyretin	1qab	A	A	2.9	51097	89910	21010	1.4E+03
retinol binding protein	1qab	B	BCD	2.9	39710	70460	19520	1.5E+03

	pdbid			Circularity	RMSD to plane	Atom burial	Atom access-ibility
Miscellaneous							
cyclophilin A	1ak4	A	A	4.0E-01	2.5E+00	6.8E+00	1.2E+01
HIV-1 capsid	1ak4	B	D	1.9E-01	1.7E+00	1.1E+01	1.8E+01
nucleotide exchangefactor GRPE	1dkg	A	A B	2.5E-01	2.5E+00	8.9E+00	1.8E+01
molecular chaperone DNAK	1dkg	B	D	2.1E-01	2.5E+00	9.2E+00	1.7E+01
Fab IgG1	1igc	A	L H	7.4E-01	1.7E+00	9.0E+00	1.6E+01
protein G	1igc	B	A	7.7E-01	1.7E+00	9.7E+00	1.6E+01
transthyretin	1qab	A	A	1.7E-01	6.5E+00	8.3E+00	1.5E+01
retinol binding protein	1qab	B	BCD	9.4E-02	5.9E+00	9.3E+00	1.5E+01

	pdbid			Shape comple-mentarity	Vol. Ratio infc core/ prot. core	Vol. Ratio infc perim/ prot. core
Miscellaneous						
cyclophilin A	1ak4	A	A	8.1E-01	1.1E+00	1.7E+00
HIV-1 capsid	1ak4	B	D	8.1E-01	9.5E-01	1.9E+00
nucleotide exchangefactor GRPE	1dkg	A	A B	6.2E-01	1.1E+00	3.0E+00
molecular chaperone DNAK	1dkg	B	D	6.2E-01	1.0E+00	2.2E+00
Fab IgG1	1igc	A	L H	6.9E-01	1.2E+00	2.0E+00
protein G	1igc	B	A	6.9E-01	1.2E+00	2.0E+00
transthyretin	1qab	A	A	5.3E-01	1.4E+00	2.1E+00
retinol binding protein	1qab	B	BCD	5.3E-01	1.0E+00	2.2E+00

	pdbid			Hbonds <3.3 A	per 100 sq angstro	Hbonds <4.0 A	per 100 sq angstro
Miscellaneous							
cyclophilin A	1ak4	A	A	1.5E+01	3.2E+00	2.1E+01	4.5E+00
HIV-1 capsid	1ak4	B	D	1.5E+01	2.7E+00	2.1E+01	3.8E+00
nucleotide exchangefactor GRPE	1dkg	A	A B	1.0E+00	1.0E-01	1.0E+00	1.0E-01
molecular chaperone DNAK	1dkg	B	D	1.0E+00	1.0E-01	1.0E+00	1.0E-01
Fab IgG1	1igc	A	L H	9.0E+00	1.4E+00	1.4E+01	2.2E+00
protein G	1igc	B	A	9.0E+00	1.3E+00	1.4E+01	2.1E+00
transthyretin	1qab	A	A	1.3E+01	9.2E-01	2.2E+01	1.6E+00
retinol binding protein	1qab	B	BCD	1.3E+01	8.7E-01	2.2E+01	1.5E+00

Miscellaneous	pdbid			%non-pola	%polar	%charged
cyclophilin A	1ak4	A	A	5.7E-01	2.8E-01	1.5E-01
HIV-1 capsid	1ak4	B	D	7.2E-01	2.8E-01	0.0E+00
nucleotide exchangefactor GRPE	1dkg	A	A B	6.4E-01	1.7E-01	1.9E-01
molecular chaperone DNAK	1dkg	B	D	6.3E-01	1.8E-01	1.9E-01
Fab IgG1	1igc	A	L H	6.2E-01	2.9E-01	9.1E-02
protein G	1igc	B	A	4.8E-01	4.7E-01	5.0E-02
transthyretin	1qab	A	A	5.6E-01	3.3E-01	1.1E-01
retinol binding protein	1qab	B	BCD	7.2E-01	1.9E-01	8.7E-02