2010

Computational Prediction and Analysis of the NAPP – DR6 Interaction: Implications for Alzheimer's Research

Joseph Audie

Sergei Y. Ponomarev

Computational Prediction and Analysis of the NAPP – DR6 Interaction: Implications for Alzheimer's Research

Sergei Y. Ponomarev[1] and Joseph Audie[2,3,4]

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that involves a devastating clinical course and lacks an effective treatment. A biochemical model for neuronal development, recently proposed by Nikolaev and co-workers, hinges on a novel protein-protein interaction between the death cell receptor six (DR6) ectodomain and an N-terminal fragment of amyloid precursor protein (NAPP). The model provides a coherent and satisfying framework for better understanding AD pathophysiology. Moreover, the DR6-NAPP interaction offers a tempting target for novel pharmacological intervention. Given all of this, we constructed a structural model of the DR6-NAPP interaction using the neurotrophin p75 receptor as a template. Importantly, crucial steps in the modeling pipeline were independently validated using the p75 receptor. The final docked model shows excellent agreement with a variety of biophysical and theoretical data sets. Particularly worth noting is the excellent observed agreement between the theoretically calculated DR6-NAPP binding free energy and the corresponding experimental quantity. In brief, the balance of the evidence suggests that the DR6-NAPP model proposed here should prove useful in future studies, modeling work and efforts aimed at structure-based drug design.

[1]Worcester Polytechnic Institute, Department of Chemistry and Biochemistry, 65 Prescott Drive, Worcester, Massachusetts 01609, *email:* sponomarev@wpi.edu

[2]Sacred Heart University, Department of Chemistry, 5151 Park Ave, Fairfield, Connecticut 06825, *email:* audiej@sacredheart.edu

[3]CMD Bioscience, LLC, 554 Boston Post Rd #318, Orange, Connecticut 06477, *email:* joseph.audie@cmdbioscience.com

[4]Correspondence should be addressed to J.A. *email:* audiej@sacredheart.edu

Introduction

Alzheimer's disease (AD) is the sixth leading cause of death in the United States. Currently, an estimated 5.3 million Americans of all ages have the disease[1]. Alzheimer's disease is a neurodegenerative disorder that is pathologically defined by the presence of amyloid depositions and neurofibrillary tangles. The AD brain is further characterized by neuronal loss, primarily in the cerebral cortex and hippocampus, and cholinergic depletion. Clinically, AD follows a progressive course marked by cognitive impairment and memory loss[2-4]. Despite several decades of research, a cure for Alzheimer's disease does not exist; available treatments offer relatively small symptomatic benefit and remain palliative in nature.

Most experts agree that the amyloid precursor protein (APP) and APP proteolysis play a role in AD. Similarly, the evidence seems to implicate neuronal loss, neuronal degeneration and axonal degeneration in the progression of disease. The available evidence also indicates a role for apoptosis in the pathophysiology of Alzheimer's disease and for pro-apoptotic caspase enzymes in particular. Indeed, it has been hypothesized that caspase-6 plays a key role in disease propagation[2,4-8].

A growing body of evidence indicates an important role for trophic factor deprivation, specifically nerve growth factor (NGF) and brain-derived nerve growth factor (BDNF) deprivation, in the etiology and pathophysiology of Alzheimer's disease. A particularly striking example of this includes the development of a mouse model (AD11 anti-NGF mice) that expresses transgenic NGF antibodies and that is marked by progressive neurodegeneration, amyloid deposition, neurofibrillary tangles, cholinergic depletion, and behavioral deficits. The role of NGF deprivation is further underscored by the successful use NGF supplementation to rescue AD11 mice and by a recent study which suggests the efficacy of NGF therapy for reversing neuronal cholinergic decline in aged, non-human primates[2,9-11].

Clearly, the elucidation of a biochemical mechanism that connects all of the data, especially trophic factor deprivation, altered APP metabolism, caspase activation, neuronal apoptosis and neuritic degeneration, would help to improve our understanding of the cellular events that underlie AD and suggest new and promising routes for novel therapeutic development. Hints of such a biochemical mechanism come from a recent study that correlates NGF and BDNF withdrawal, neurotrophin receptor p75 activity, amyloidogenesis, and apoptosis in hippocampal neurons[12]. A still more recent study, published by Nikolaev and co-workers, suggests a new

biochemical model for AD that accounts for the data of the previously mentioned study and that also includes an explicit role for caspase-3 and caspases-6 in mediating apoptotic neuronal degeneration and axonal degeneration, respectively. The model suggested by Nikolaev and co-workers, however, introduces a new protein into the story termed death cell receptor 6 (DR6), a trasmembrane receptor with an intracellular death domain that appears to be activated by an N-terminal APP fragment (NAPP)[13].

The primary focus of the Nikolaev et al. paper was the elucidation of a novel molecular mechanism to account for axonal pruning and neuronal cell death during physiological development. According to the developmental model proposed by Nikolaev et al, tropic factor deprivation results in APP proteolysis, culminating in the release of an N-terminal APP fragment (NAPP) into the extracellular milieu. NAPP then serves as a ligand for DR6 and activates DR6. Activation of DR6 results in the subsequent downstream activation of caspase3 and caspase6, respectively, resulting in accelerated neuronal apoptosis, neuronal degeneration, axonal degeneration and the physiological sculpting of nerve connections in the developing brain. Nikolaev and co-workers suggested that this physiological pathway could be hijacked in the adult brain, possibly resulting in AD.

Hence, the AD model proposed by Nikolaev and co-workers explicitly links trophic factor deprivation with neuronal loss and degeneration through a series of well characterized biomolecular events that might be susceptible to therapeutic intervention. In particular, the model proposed by Nikolaev and co-workers suggests an obvious opportunity for innovative therapeutic intervention in the form of rationally designed molecular inhibitors of the NAPP-DR6 protein-protein interaction. Such a therapeutic approach could be used as an alternative or supplement to NGF therapy and other therapeutic modalities. Moreover, the NAPP-DR6 interaction has the advantage of being an upstream event and there is growing interest in the rational design of therapeutic protein-protein modulators[14].

The rational design of molecular antagonists of the DR6-NAPP interaction would clearly benefit from an atomic-level or residue-level structural model of the DR6-NAPP complex. Similarly, our collective theoretical understanding of the interaction and our ability to design and interpret experiments would also benefit from a structural model. Unfortunately, such a molecular model is lacking. Indeed, while a monomeric crystal structure of the growth factor like domain of NAPP is available (GFD NAPP), a monomeric structure of the DR6 ectodomain

is not[15]. Importantly, Nikolaev and co-workers performed an experiment using an antibody specific for the growth factor like domain (GFD) of NAPP (anti-GFD NAPP (22C11)) that blocked DR6 binding, a result which suggests a key role for GFD mediated binding to DR6.

Given all of this, we hypothesized that GFD NAPP forms a stable interaction with DR6 and labored to produce a theoretical model of the DR6- GFD NAPP interaction through the use of advanced homology modeling, docking, energy minimization and binding free energy calculations. More specifically, we used I-TASSER and implicit solvent energy minimization to build and refine a homology model of the DR6 ectodomain[16-22]. We then used the ClusPro rigid-body docking server to dock the DR6 model to the crystal structure of GFD NAPP[23,24]. This was followed by rigid-body energy minimization, implicit solvent pseudo-free energy scoring and final scoring and ranking using a recently described empirical free energy function[25-27]. The final model was verified against the available biophysical data, including a novel theoretical data set that was generated using a second and independent computational procedure for predicting protein-protein interface residues. Perhaps most importantly, the theoretical binding affinity implied by our DR6-GFD NAPP model is shown to be in quantitative agreement with the experimentally determined binding affinity measured by Nikolaev and co-workers. To the best of our knowledge, such quantitative free energy testing represents a milestone in theoretical model construction and testing. Based on extensive testing, we are confident that the rigid-body binding geometry of our model is roughly accurate and that many of the interface residues predicted by our model are the interface residues that will one day be revealed by experiment, especially with regards to the GFD NAPP interface residues. In sum, the available data suggests that our DR6-GFD NAPP structural interaction model is, at least at the residue-level, a reasonably good model that should prove useful in structure-based antagonist design and in the design and analysis of experiments.

Results

The primary aim of the present study was to construct a theoretical model of the DR6-GFD NAPP interaction, an interaction that may have implications for AD. A DR6-GFD NAPP interaction model was constructed using homology modeling, rigid-body docking, energy minimization, MM-GB/SA pseudo free energy scoring and empirical free energy scoring. At each step in the computational pipeline, theoretical calculations were compared, to the extent

permitted by the available data, with experimental results and independently generated theoretical results.

Modeling and evaluating our DR6 homology model and the GFD NAPP crystal structure. Figure 1 provides a ribbon representation of the GFD NAPP crystal structure (28-123). GFD NAPP is a high resolution (1.8 Å), high quality (R-value = 0.203; R-free value = 0.242) crystal structure that exhibits a globular fold. Several key residues (66-81) that comprise a lone alpha-helix-loop motif are highlighted in cyan. Anti-GFD NAPP antibody 22C11 interacts with the alpha-helix-loop motif and blocks DR6-GFD NAPP binding.

Figure 2 provides a ribbon representation of our refined DR6 ectodomain homology model. The ectodomain of DR6 comprises residues 67-211. The DR6 ectodomain model takes on a more extended shape and exhibits beta secondary structure interconnected through less well defined structural elements. The DR6 structure appears to form a structural depression or basket region that seems well suited to accommodate a globular protein such as GFD NAPP. The DR6 homology model was evaluated using a variety of computational tools and the results are summarized in Table 1. The DR6 homology model was ultimately docked to the GFD NAPP crystal structure.

Modeling the interaction between DR6 homology and GFD NAPP. Figure 3 provides a ribbon representation of the final or best DR6-GFD NAPP interaction model. The model indicates an important recognition role for the GFD NAPP alpha-helix-loop motif (residues 66-81). It also appears that the GFD NAPP alpha helix rests in or lines the previously mentioned DR6 structural depression or basket. Worth noting is that the empirically predicted binding affinity implied by the docked model (-11.1 kcal/mol) is in excellent agreement with the experimentally estimated binding free energy (-11.5 kcal/mol).

Biophysical testing of the DR6-GFD NAPP interaction model. Both NAPP and an N-terminal fragment of APLP2 have been shown to bind the DR6 ectodomain. Given this, we judged it worthwhile to align the sequences between GFD NAPP and APLP2. Figure 4 displays the sequence alignment. Secondary structural information is also presented. The anti-GFD NAPP (22C11) antibody binding epitope is indicated by a solid black line.

The GFD NAPP binding site residues suggested by our docking study were compared to the GFD NAPP binding site residues suggested by the experimentally observed inhibitory activity of the anti-GFD NAPP 22C11 antibody. The results are graphically summarized in Figure 5.

Nikolaev and co-workers showed that the neurotrophin receptor p75 binds NAPP, albeit with lower affinity ($EC_{50} \approx 300$ nM) than does DR6. Given this, Figure 6 provides the results of a pairwise sequence alignment between the DR6 ectodomain and p75. Also provided in Figure 6 is secondary structural information for the DR6 ectodomain.

Theoretical testing of the DR6-GFD NAPP interaction model. PPI-Pred was used to predict binding patches for GFD NAPP and the ectodomain of DR6. For GFD NAPP, PPI-Pred produced two patches (I and II). For DR6, PPI-Pred calculated three binding patches (I, II, and III). The results for GFD NAPP are graphically summarized in Figure 5. Running from left to right (a, b, and c) are structural models that depict (in cyan) interface residues derived from the 22C11 antibody binding experiments, putative GFD NAPP interface residues derived from the present docking study, and potential interface residues obtained from the PPI-Pred calculations. Only the calculated PPI-Pred residues that agree with the residues obtained from docking study are shown. The results are also summarized in Table 2.

As in the case of GFD NAPP, putative DR6 interface residues derived from the docking study were compared to potential DR6 interface residues derived from PPI-Pred. Unlike the case with GFD NAPP, an experimentally derived DR6 interface residue set proved to be unavailable. Only the calculated PPI-Pred residues that agree with the residues obtained from docking are shown. The results are summarized in Figure 7 and in Table 3.

Discussion

A growing body of evidence implicates trophic factor deprivation, APP metabolism, pro-apoptotic caspase activity, neuronal apoptosis, and neuronal and neuritic degeneration in AD.
In a recent article, Nikolaev and colleagues describe a novel interaction between the death cell receptor six ectodomain (DR6) and an N-terminal fragment of APP (NAPP). While their work was primarily focused on understanding the biochemical basis of neural development, they nonetheless considered the possible implications of their work for AD. Indeed, the DR6-NAPP interaction elucidated by Nikolaev et al. might provide the lynchpin for constructing a coherent

and satisfying framework for understanding the pathological relationship between trophic factor deprivation, APP processing, caspase mediated neuronal apoptosis and degeneration, and AD disease progression.

According to the developmental model of Nikolaev and colleagues, following trophic factor deprivation APP is proteolytically cleaved, resulting in the release of N-terminal amyloid precursor protein (NAPP) into the extracellular space. This is thought to be followed by NAPP binding to the death cell receptor six (DR6) ectodomain (67-211), resulting in neuronal apoptosis and neuronal and axonal degeneration through the downstream mediation of caspase-3 and caspase-6. It is through this mechanism that neuronal connections are pruned and sculpted during development. It is speculated that this pathway is high jacked in the AD brain.

Despite its obvious importance, the structural details that characterize the DR6-NAPP interactions are unclear. In an effort to elucidate some of those structural details, we labored to produce a theoretical model of the DR6-NAPP interaction through the use of homology modeling, rigid-body docking, implicit solvent energy optimization and theoretical free energy scoring. Such a structural model would greatly aid with the design and interpretation of experiments and with future structure-based drug discovery efforts.

Nikolaev et al. performed binding experiments which indicate that NAPP binds DR6 with low nanomolar affinity ($EC_{50} \approx 4.6$ nM or $\Delta G_{bind,exp} \approx -11.5$ kcal/mol). Additional biophysical evidence presented by Nikolaev et al. strongly suggests that the growth factor domain (GFD) of NAPP (GFD NAPP) is the primary DR6 ectodomain interaction partner. In particular, the fact that the GFD NAPP antibody 22C11 blocks NAPP binding to DR6 provides near conclusive evidence that GFD NAPP provides the primary or even exclusive interaction surface for the DR6-NAPP interaction. Importantly, a high resolution crystal structure of GFD NAPP (28-123) is available. Unfortunately, a DR6 ectodomain structure is lacking. Nevertheless, we were able to make progress towards our ultimate goal by constructing an optimized DR6 homology model. The optimized DR6 ectodomain model was subsequently docked to the GFD NAPP crystal structure and the best docked model was selected according to an empirical estimate of the binding affinity. The best docked DR6-GFD NAPP model was then compared against the available biophysical data and independently generated theoretical data. The balance of the evidence suggests that our DR6-GFD NAPP model captures the essential structural details of DR6 binding to NAPP, especially with regard to the residue-level contribution of GFD NAPP.

The GFD NAPP crystal structure and the accuracy of our DR6 homology model. The primary goal of the present study was to construct a structural model of the interaction between the DR6 ectodomain and the growth factor-like domain (GFD) of NAPP. Our first task, then, was to obtain structural models of the DR6 ectodomain and GFD NAPP, respectively. Our second task was to verify the physical soundness and accuracy of the structures.

We were able to locate a high resolution (1.8 Å) crystal structure of GFD NAPP that included residues 23-128. The excellent R-value (0.203), R-free value (0.242) and unblemished Ramachandran plot (data not shown) indicate an excellent structure that is eminently suited for modeling work. As indicated in Figure 1, the GFD NAPP structure reveals a predominantly beta globular structure with a single, prominent alpha helical-loop segment that spans residues 66-81. As discussed below, this alpha helix-loop motif appears to play an important role in GFD NAPP-DR6 binding and recognition.

An experimentally determined DR6 structure is currently unavailable. Thus, we decided to construct a homology model of the DR6 ectodomain using residues 67-211 and the I-TASSER modeling server (Figure 2). Multiple lines of evidence suggest that our I-TASSER DR6 ectodomain homology model is a reasonably accurate one that is suitable for modeling work.

First, the neurotrophin receptor p75 (PDB code: 1sg1, X chain) that served as the template for our DR6 ectodomain model is a good template structure. The 1sg1 crystal structure is a reasonably high quality crystal structure with good resolution (2.40 Å) and an acceptable R-value (0.245) and R-free value (0.269). With the exception of Asp112, all non-glycine residues are located within the allowed regions of the Ramachandran map (data not shown). Like DR6, the p75 receptor is a member of the tumor necrosis factor (TNF) family. Importantly, the 1sg1 crystal structure provides the coordinates of the p75 receptor in complex with its nerve growth factor (NGF) ligand and thus represents a binding competent conformation. Moreover, Nikolaev et al. showed that the p75 receptor was the only other TNF receptor tested that formed a stable interaction with NAPP ($EC_{50} \approx 300$ nM). Finally, the ≈ 28% sequence identity between the entire DR6 target sequence and the p75 receptor sequence is good and the overall sequence coverage includes the entire DR6 ectodomain (67-211). These structural, biophysical and functional considerations strongly suggest that p75 is a good template for constructing a high quality structural model of the binding competent conformation of the DR6 ectodomain.

Second, we built our DR6 model using the I-TASSER server, a homology modeling server that has been widely employed by the scientific community and that was ranked first in both the CASP7 and CASP8 experiments. The I-TASSER server follows an especially demanding computational workflow. In particular, I-TASSER automatically selects appropriate protein template structures and builds a model of the target or query structure (DR6) through a rigorous procedure that involves fragment reassembly of aligned regions, *ab initio* model construction of unaligned regions, clustering, energy scoring and optimization of the target structures hydrogen bonding network. These methodological considerations further suggest the physical accuracy and reliability of our DR6 ectodomain model.

Third, the internal quality checks automatically calculated by the I-TASSER server further suggest that our DR6 ectodomain model is an accurate one. In particular, for the top ranked model the I-TASSER server calculates a C-score, a predicted TM-score and a predicted RMSD value. Briefly, the C-score represents a confidence score for estimating the overall quality of a model. Typically, C-score values are in the range of -5 to 2, with a higher value indicative of a model with high confidence. The C-score of our DR6 model is on the higher side of that range (1.31) which suggests a high confidence model. The target TM and RMSD values provide quantitative estimates of how much the I-TASSER homology model is predicted to structurally deviate from a hypothetical crystal structure. A TM score > 0.5 indicates a model of correct topology and an RMSD < 0.9 would indicate a model that is predicted to be within the unavoidable error of a crystal structure. The TM-score (0.9±0.06) and RMSD value (2.2±1.7) calculated for our DR6 model thus indicate a very good model that is well-suited for modeling work. The results of the I-TASSER quality analysis are summarized in Table 1.

Fourth, in addition to the internal I-TASSER quality checks we calculated additional quality checks using three independent server-based methods (ProSA, Qmean, and DFIRE). All three methods provide independent verification that our DR6 homology model is a reasonably accurate one. The ProSA server calculates an overall z-score that can be used to evaluate the structural quality of a computational protein model with respect to crystal structures and NMR structures. The Qmean server employs a five-term composite scoring function that includes a torsional potential, a pairwise interaction potential, a solvation term, and two terms that quantify the level of agreement between predicted secondary structure and residue accessibilities and observed model secondary structure and residue accessibilities. The Qmean server outputs a pseudo

energy that can be used to asses model quality; the lower the calculated Qmean score the better. DFIRE is an all-atom statistical potential that can be used to assess non-bonded interactions in a given protein model. A pseudo energy for the entire model is provided which reflects the quality of the model. A lower energy indicates a more native-like model. The results from the ProSA, Qmean and DFIRE calculations are provided in Table 1 and clearly indicate that our DR6 ectodomain model is reasonably accurate and useful for further modeling work and analysis.

Finally, we attempted to refine and improve our DR6 model through energy minimization with an implicit solvent model. The problem of refining reasonably good homology models is a difficult but important one. Recent work by Summa, Chopra and Levitt, suggests that energy minimization using an all-atom molecular mechanics force field or knowledge-based potential, especially when combined with an implicit (GB/SA) solvent model, can be used to improve the structural agreement between model structures and crystal structures. In particular, their work suggests that energy minimization combined with an implicit treatment of the solvent can be used to hold crystal structures in their native conformations and move decoy structures into better agreement with crystallographic ones[19-21]. Given this, we energy minimized our DR6 ectodomain homology model using the Amber99 force field and GB/SA implicit solvent model, as implemented in the TINKER molecular modeling package. Against the backdrop of the results of Summa, Chopra and Levitt, the successful convergence of the minimization and the extremely low relative conformational free energy (-4564.2 kJ) suggests that the minimization resulted in a superior DR6 model. The results are summarized in Figure 2 and Table 1. The energy minimized DR6 model served as the input receptor structure we used in the DR6-GFD NAPP rigid-body docking study to be discussed below.

Building the DR6-GFD NAPP complex model. The successful construction of a DR6 homology model was a means to an end, for our primary goal was to model the interaction between the DR6 ectodomain and GFD NAPP. To do this, we employed rigid-body computational docking using the ClusPro server. The ClusPro docking server was selected because of the long and rigorous publication history that underlies its basic methodology, its widespread use by structural biologists, and its impressive record of success in the various CAPRI competitions[28,29].

Following the ClusPro calculations, the top 10 docked models were subject to rigid-body energy minimization and scored and ranked according to a pseudo binding free energy that was

calculated using a molecular mechanics potential and GB/SA implicit solvent model. Docked complexes with a positive calculated pseudo binding free energy were filtered out and eliminated from further consideration. Work in our lab and by other groups suggests that this protocol or similar protocols can result in more accurate complex structures and an enhanced ability to identify accurate or native-like protein-protein interactions (our data is unpublished)[30,31].

The binding affinities of the selected docked models were then estimated using a recently described empirical free energy function[25-27]. The empirical free energy function we employed estimates the binding affinity according to a regression-weighted sum of six physical descriptors. The six descriptors include the number of hydrophobic and charged groups that are buried on complex formation, the number of buried interface salt bridge and hydrogen bonding interactions, the number of side chain torsions that are immobilized on binding and the interface void or gap volume. The function has been exhaustively tested and has been shown to have an accuracy of ≈ 1.0 kcal/mol and to be capable of correctly ranking native and native-like binding modes with respect to non-native ones. This hierarchical free energy estimation protocol ultimately converged on a single DR6-GFD complex model that was selected as the best model of the DR6-GFD NAPP interaction. The fundamental propriety of our docking procedure is strongly suggested by prior validation on the neurotrophin receptor p75-NGF interaction (see Supplementary information Figures 1-3, Tables 1-4).

Testing of the DR6-GFD NAPP complex model. *A priori* considerations regarding the rigorous and systematic nature of our computational modeling and evaluative protocol suggested that our final DR6-GFD NAPP model would be a physically reasonable one. The next step was to test the model *a posteriori*. In what follows, we divide the testing into two categories: (1) biophysical testing of the docked DR6-GFD NAPP model, and (2) theoretical testing of the docked model. The biophysical model testing phase involved binding affinity comparisons, the analysis of GFD NAPP and DR6 sequence alignments and a comparative analysis of anti-GFD NAPP 22C11 antibody data. The theoretical model testing phase involved comparisons between data derived from our DR6-GFD NAPP model and independently generated computational data.

Biophysical testing of the DR6-GFD NAPP complex model. An experimentally-based estimate of the of binding affinity for the DR6-GFD NAPP interaction agrees almost perfectly with a

theoretical estimate of the binding affinity derived from our DR6-GFD NAPP model (Figure 3). This counts as impressive evidence that our model interface is a reasonable one. In particular, Nikolaev at al. used an ELISA based technique to experimentally determine the binding affinity of the DR6-NAPP interaction ($EC_{50} \approx 4.6$ nM). This provided us with the opportunity to compare the empirically predicted binding affinity implied by our DR6-GFD NAPP model (Eq. (1), ΔG_{bind} = -11.1 kcal/mol) with the corresponding experimental quantity ($\Delta G_{bind,exp} \approx$ -11.5 kcal/mol). The observed level of agreement between theory and experiment is excellent and while it doesn't prove that our model is an accurate one, it nevertheless provides strong verification of our model. The results of this test are all the more impressive given that the same empirical free energy function was used to accurately calculate the binding affinity of the p75-NGF interaction (-11.8 kcal/mol predicted / -12.4 kcal/mol experimental) that served as the template for our DR6 homology model (see Supplementary information Table 4).

In addition to binding NAPP, Nikolaev et al. also showed that the DR6 ectodomain binds the N-terminus of APLP2. Thus, we compared the predicted interface residues of GFD NAPP, derived from our DR6-GFD NAPP model, with the aligned residue positions of APLP2. On the assumption that interface residues tend to be conserved, we reasoned that if the predicted GFD NAPP interface residues were accurate they would display good agreement with the aligned APLP2 residues. In summary, the GFD NAPP interface residues predicted by our docked model align nicely with the APLP2 residues that probably mediate binding between APLP2 and DR6 (Figure 4). This provides indirect evidence that the residue-level contribution of GFD NAPP to DR6 binding is captured by our model.

Nikolaev et al. also showed that the NAPP antibody 22C11 interferes with DR6-NAPP binding. Importantly, Hilbich et al. showed that the binding epitope recognized by the 22C11 antibody spans NAPP residues 66-81. This represents a stretch of residues that are localized around the lone helix (66-76) of GFD NAPP. On the assumption that 22C11 blocks DR6 binding to NAPP by binding to the same GFD NAPP surface that mediates DR6-GFD NAPP binding, we compared the GFD NAPP interface residues derived from our model with the GFD NAPP epitope that is known to bind 22C11. Once again, we reasoned that good agreement between the two would tend to verify our model. Using a 4.5 Å cutoff criterion, the GFD NAPP residues that line the DR6-GFD NAPP interface of our model include residues 67, 68, 70, 71, 74, 78, and 79. There thus exists excellent agreement between the experimentally determined 22C11

epitope and the interface residues of our model. Thus, the modeled GFD NAPP contribution to DR6 binding enjoys further verification and, moreover, focuses attention on the specific role played by helix residues 66-76 in DR6-NAPP recognition.

In addition to binding DR6, Nikolaev and coworkers also showed that NAPP binds the p75 receptor, albeit with reduced affinity. Thus, we attempted to align the p75 sequence with the DR6 ectodomain sequence on the assumption that the DR6 and p75 NAPP binding site residues are conserved. Unlike the case with APLP2, however, the DR6 and p75 alignment results, as summarized in Figure 6, paint an ambiguous and difficult to interpret picture. Out of the 24 residues (96, 98-99, 103-104, 107-108, 110, 120-121, 128, 130-132, 139, 141-142, 158-159, 162-165, 167) that comprise the DR6 contribution to our model interface, only 5 (21%) are identical to the corresponding residues in p75. This is slightly less than the overall level of conservation that characterized the alignment between p75 and DR6 that was used to construct the I-TASSER DR6 homology model. If we include similar residues, the percentage of DR6-p75 matches increases to 37.5%. Interpretation of the results is further complicated by various nonconservative substitutions. For example, p75 substitutes an Arg for DR6 Val167 and a second Arg for DR6 Asp98. Such nonconservative substitutions could explain the reduced affinity of p75 for NAPP. These results suggest that the DR6 binding motif is either not strongly conserved by p75 or that our DR6 interface residue model predictions are inaccurate. We think the later possibility is unlikely given all of the evidence presented, the relatively weak binding of p75, and given that high sequence conservation alone is an imperfect predictor of shared binding motifs[32].

Theoretical testing of the DR6-GFD NAPP complex model. Typically, theoretical protein-protein or protein-ligand structural models are tested against experimental mutation data[33]. Unfortunately, we were not able to locate any such data for the DR6-GFD NAPP interaction. Despite this, we were able to generate similar data using an independent computational procedure. Briefly, we used the protein-protein interaction prediction server or PPI-Pred to predict binding surfaces and thus potential interface residues for DR6 and GFD NAPP, respectively. We then compared the set of predicted interface residues, generated using PPI-Pred, with the set of interface residues derived from our model. We reasoned that good agreement between the two independently generated data sets would tend to confirm the basic

accuracy of our model (Figure 5, Figure 7, Table 2 and Table 3). Importantly, we first validated the procedure using the known structural interaction between the neurotrophin p75 receptor and its nerve growth factor (NGF) ligand (see Supplementary information Figures 1-3, Tables 1-4).

In the case of GFD NAPP, PPI-Pred predicated two possible binding patches (I and II); in the case of DR6, PPI-Pred predicted three possible binding patches (I, II, and III). Out of 44 PPI-Pred predicted residues for GFD NAPP, 15 agree perfectly with interface residues derived from the GFD NAPP binding interaction that characterizes our docked model. The PPI-Pred patch I and III predictions for DR6, failed to show any agreement with our docked model. The 31 PPI-Pred patch II residue predictions, however, produced 8 residue hits when compared to the DR6 residues that contribute to the interface of our docked model. Hence, many of the DR6 and GFD NAPP interface residues that characterize our docked model are also independently predicted to be interface residues. This provides still more evidence that our DR6-GFD NAPP interaction model, especially regarding the GFD NAPP contribution, is an accurate one.

Conclusion

AD is a devastating disease with no effective treatments. Nikolaev and co-workers have proposed a biochemical model of AD that plausibly accounts for much of the available data. Key to their model is a novel interaction between the death cell receptor six (DR6) ectodomain and an N-terminal fragment of amyloid precursor protein (NAPP), with particular emphasis on the growth factor like domain of NAPP (GFD NAPP). The elucidation of the structural basis of the DR6-GFD NAPP interaction could pave the way towards improved understanding of AD and innovative pharmacological therapies for AD. Here we describe a computational work flow for constructing a structural model of the DR6-GFD NAPP complex. The work flow adopted combined and integrated homology modeling, quantitative homology model testing and evaluation, implicit solvent homology model energy minimization, rigid body protein-protein docking and refinement, and a hierarchical method of protein-protein binding free energy scoring and ranking. The use of such a comprehensive modeling pipeline strongly suggested that the final DR6-GFD NAPP model would be a reasonable one. Indeed, extensive testing against multiple lines of theoretical and experimental evidence strongly suggests that the DR6-GFD NAPP model described here is a good one, especially with regards to the residue-level contribution of GFD NAPP. Interestingly, we show that the theoretically predicted binding

affinity for the DR6-GFD NAPP interaction is in quantitative agreement with the experimentally determined binding affinity. To the best of our knowledge, this represents a first in the blind modeling of a novel protein-protein interaction. In summary, the DR6-GFD NAPP interaction model proposed here satisfactorily accounts for a wide range of disparate evidences and as such, should prove helpful in the design and interpretation of *in vitro* and *in silico* experiments aimed at understanding and modifying the DR6-GFD NAPP and its connection to AD.

Acknowledgments

We would like to thank Dr. Anatoly Nikolaev, Genentech Inc., for the valuable discussions and suggestions on the topic of AD.

Author Contributions

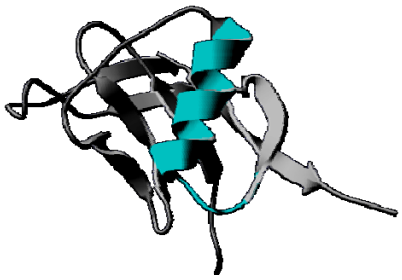
Both authors contributed equally to the manuscript. J.A. supervised the project, predicted the binding affinities (scoring function), prepared the tables and some figures, edited the supplementary materials and wrote the paper. S.Y.P. proposed the basic research idea, performed most of the molecular modeling procedures (homology modeling, docking, minimization), prepared majority of the figures, wrote the supplementary materials, and proof read the manuscript.

Competing Financial Interests

The authors declare no competing financial interests. J.A. is a co-founder of CMDBioscience.

Figure Legends and Tables

Figure 1: Ribbon representation of the GFD NAPP crystal structure with key residues highlighted in cyan.

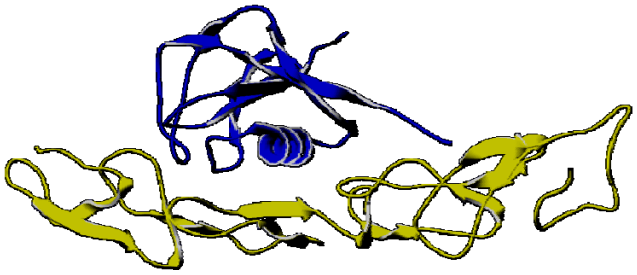


The PDB identifier for the structure is 1mwp. The structure comprises residues 28-123 and was solved at 1.8 Å resolution; the structure has an R-value = 0.203 and an R-free value = 0.242. The cyan colored alpha helix and loop span residues 66-81 and makes up the known binding epitope for the anti-GFD NAPP antibody 22C11. Importantly, the 22C11 antibody is known to interfere with DR6-GFD NAPP binding. Hence, it is reasonable to assume that the same helix-loop motif makes up a key recognition element for DR6-GFD NAPP binding. Importantly, according to our best docked model GFD NAPP residues E67, G68, L70, Q71, Q74, P78, and E79 are all predicted to be interface residues.

Figure 2: Ribbon representation of the refined DR6 ectodomain homology model.

The refined DR6 ectodomain (67-211) homology model was built using the I-TASSER modeling server and was subject to energy minimization using a GB/SA implicit solvent model.

Figure 3: Ribbon representation of and predicted binding affinity for the best DR6-GFD NAPP interaction model.



a

Complex	Predicted ΔG (kcal/mol)	Experimental ΔG (kcal/mol)
DR6 - GFD NAPP	-11.1	-11.5

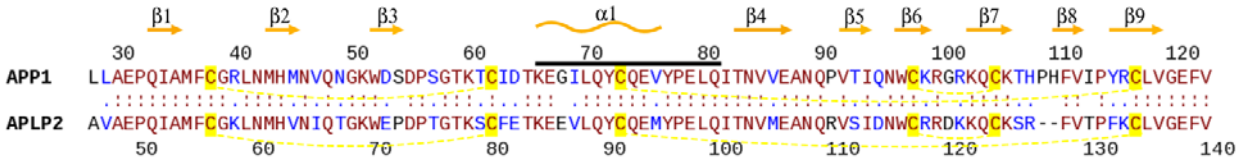
b

a) The refined DR6 ectodomain (67-211) homology model was docked to the GFD NAPP crystal structure using the rigid-body ClusPro protein-protein docking server. The model clearly reveals an important recognition role for the lone GFD NAPP alpha helix, in good agreement with the 22C11 antibody data (see Figure 1 and 4 for more details). According to our methodology, the top 10 ClusPro models were selected for further analysis. All 10 ClusPro complexes were subject to rigid-body energy minimization and scored using a standard MM-GB/SA methodology. Complexes with negative MM-GB/SA binding free energies were then scored using an empirical free energy function.

b) The ClusPro model with the best empirical binding free energy (most negative predicted binding affinity using Eq. (1)) was selected as the best complex model and is displayed above. The predicted binding

affinity (-11.1 kcal/mol) for the best ClusPro model compares favorably with an experimental value for the DR6-NAPP binding affinity (-11.5 kcal/mol).

Figure 4: Sequence alignment and secondary structure of the growth factor-like domain of human



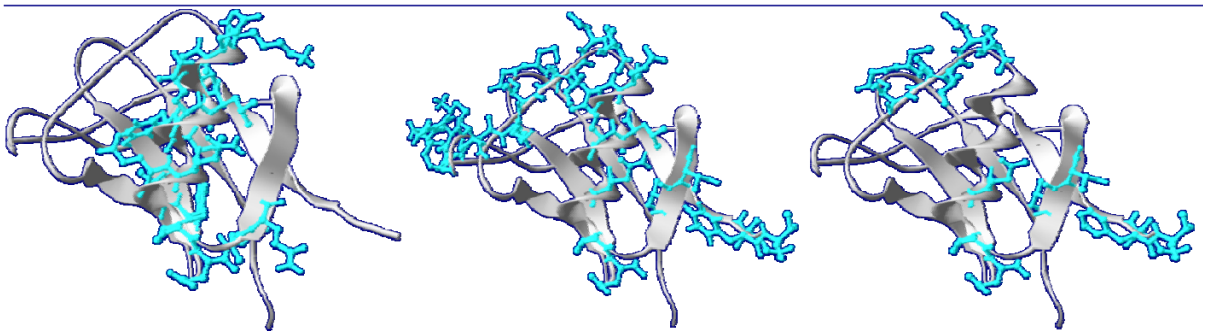
N-terminal APP (GFD NAPP) and APLP2.

Disulfide bonds are shown as yellow dash lines, identical residues are depicted in red, conserved residues are in blue, and non-conserved residues are colored black. The alignment was performed via the LALIGN server (http://www.ch.embnet.org/software/LALIGN_form.html).

Both NAPP and an N-terminal fragment of APLP2 have been shown to bind the DR6 ectodomain. Thus, given the reasonable assumption that binding motifs are highly conserved between protein ligands that bind to the same receptor, we aligned the two sequences and calculated percent identify and similarity scores. For the global alignment the percent identity and the percent similarity were calculated to be 71% and 93%, respectively. Using the ClusPro predicted interface residues (identified according to a 4.5 Å cutoff) for GFD NAPP and the corresponding aligned positions for APLP2, the percent identity and similarity were calculated to be 76% and 92%, respectively. Keeping in mind that such an analysis is plagued by uncertainty, the alignment data can be plausibly interpreted in support of our docked DR6-GFD NAPP model.

The anti-GFD NAPP (22C11) antibody binding epitope is indicated by a solid black line. Nikolaev and co-workers showed that 22C11 blocks the interaction between DR6 and GFD NAPP. Thus, it is reasonable to assume that the DR6-GFD NAPP interaction, at least in part, is mediated by some of the same residues that comprise the 22C11 binding epitope. And, this appears to be the case (see other Figures, Tables and text for further details).

Figure 5: Ribbon representations of GFD NAPP along with 22C11, ClusPro predicted, and PPI-Pred predicted interface residues (highlighted in cyan).



a **b** **c**

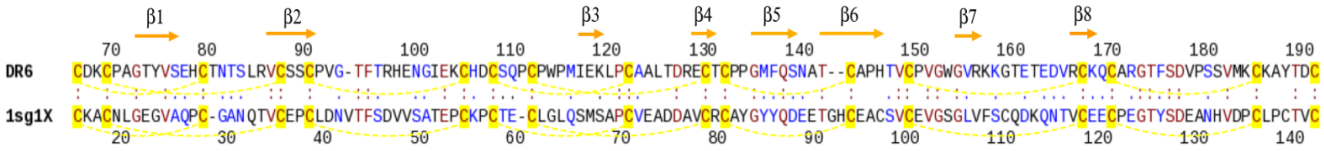
Using a crystal structure of GFD NAPP (1mwp) and a homology model of the DR6 ectodomain, we constructed a structural model of their interaction using the ClusPro docking server. We then compared the predicted ClusPro interface residues with the binding epitope of the 22C11 anti-GFD NAPP antibody that is known to block binding to the DR6 ectodomain. We also compared the ClusPro residues with interface residue predictions generated using the protein-protein interaction prediction server (PPI-Pred). In the case of GFD NAPP, PPI-Pred produced two predicted binding patches (I and II). The above displayed residues were derived from both patches; only PPI-Pred residues that agree with the ClusPro residues are displayed. We reasoned that overlap between all three disparate residue sets would tend to verify our model; this appears to be the case.

a). Ribbon representation of GFD NAPP along with the 22C11 antibody binding epitope residues shown in cyan as "balls-and-sticks".

b). Ribbon representation of GFD NAPP along with ClusPro predicted interface residues shown in cyan as "balls-and-sticks". A residue was identified as a ClusPro interface residue according to a 4.5 Å inter-chain distance criterion.

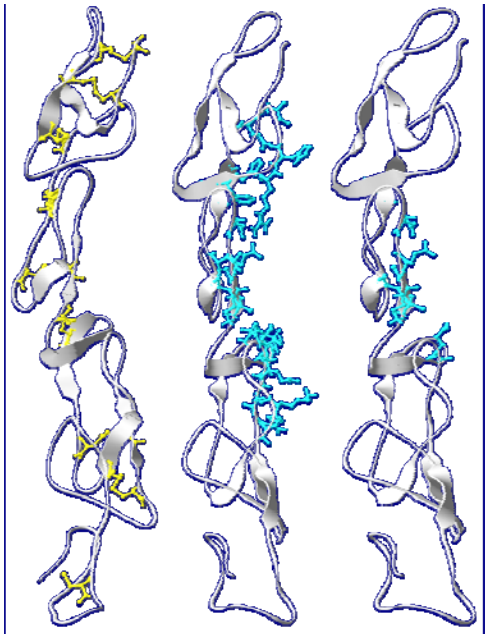
c). Ribbon representation of GFD NAPP along with PPI-Pred predicted interface residues shown in cyan as "balls-and-sticks". The PPI-Pred residues were derived from PPI-Pred predicted patches (I and II) and only residues that overlap with the ClusPro residues are displayed.

Figure 6: Sequence alignment and secondary structure of the human DR6 ectodomain and its homolog, p75 (1sg1, chain X).



The alignment was calculated using the LALIGN server. Disulfide bonds are shown as yellow dashed lines, identical residues are depicted in red, conserved residues are shown in blue, and non-conserved residues are shown in black. The basic logic for using the alignment to test our DR6-GFD NAPP protein-protein interaction model is described in the text and is summarized in Figure 1. For the global alignment, the percent identify was calculated to be 30% and the percent similarity was calculated to be 54%. For the ClusPro predicted interface residues, the values were calculated to be 27% and 36%, respectively. At best, the results of this analysis are inconclusive.

Figure 7: Ribbon representations of the DR6 ectodomain homology model along with ClusPro predicted and PPI-Pred predicted interface residues (highlighted in cyan).



a **b** **c**

See the text and Figure 2 for more details. The presented data tends to verify our DR6-GFD NAPP model.

a). Ribbon representation of the DR6 ectodomain homology model structure. The structure was built using the I-TASSER modeling server. Cysteine residues are shown as yellow "balls-and-sticks".

b). Ribbon representation of the DR6 ectodomain along with ClusPro predicted interface residues colored in cyan and depicted as "balls-and-sticks".

c). Ribbon representation of the DR6 ectodomain along with PPI-Pred predicted interface residues colored in cyan and depicted as "balls-and-sticks". The PPI-Pred residues were derived from PPI-Pred predicted patches (I, II and III) and only residues that overlap with the ClusPro residues are displayed.

Table 1: We built a DR6 ectodomain (67-211) homology model using I-TASSER. The structural and physical quality of the I-TASSER model was evaluated using a variety of internal (I-TASSER) and external (ProSA, QMEAN, and DFIRE) tools.

Model Quality	DR6-GFD NAPP I-TASSER homology model
I-TASSER C-score*	1.31
I-TASSER TM-score*	0.9
I-TASSER RMSD*	2.2
ProSA Z-Score**	-3.87
QMEAN Score***	0.445
DFIRE Energy****	-123.69
Minimization Energy*****	-4564.22

*I-TASSEER server http://zhang.bioinformatics.ku.edu/I-TASSER/

**ProSA Server https://prosa.services.came.sbg.ac.at/prosa.php

***Qmean Server swissmodel.expasy.org/qmean/

****http://swissmodel.expasy.org/workspace/

*****TINKER GB/SA AMBER99 minimization

C-score values range from -5 to 2, with a higher value indicative of a good model. A TM-score > 0.5 is associated with a model of correct overall topology. A calculated RMSD value of 2.2 suggests a model with relatively high resolution. The ProSA, Qmean and DFIRE scores all suggest a high quality model. More dtail is available in the text.

The DR6 model was refined using energy minimization along with the Amber99 force field and GB/SA implicit solvent model, as implemented in the TINKER molecular modeling package.

Table 2: Comparisons between three different methods for predicting interface residues for GFD NAPP.

ClusPro predicted interface residues for GFD NAPP:
C38,G39,T59,K60,T61,C62,I63,D64,T65,E67,G68,L70,Q71,Q74,P78,E79,I82,T83,K99,R100,K103, Q104,E121,F122,V123
PPI-Pred predicted interface residues for GFD NAPP
C38,G39,T61,C62,I63,D64,T65,Q74,P78,E79,I82,T83,E121,F122,VAL123
22C11 predicted interface residues for GFD NAPP:
K66,E67,G68,I69,L70,Q71,Y72,C73,Q74,E75,V76,Y77,P78,E79,L80,Q81

The first row provides the interface residue predictions or contributions of GFD NAPP implied by our DR6-GFD NAPP ClusPro docked model, using a 4.5 Å inter-chain cutoff criterion. The residues provided in the first row provide the basis for comparison with the bottom two rows. Residue agreement with the first row is thus indicated by underlining residues in the bottom two rows. Substantial agreement between the three independently generated data sets would tend to verify our docked model and this appears to be the case.

The second row provides interface residue predictions for GFD NAPP that were generated using PPI-Pred. Only PPI-Pred residues that agree with the ClusPro residues are underlined. For GFD NAPP, PPI-Pred predicted two binding patches (I and II). Patch I has 25 residues; 8 overlap with the ClusPro interface residues; the first 8 residues above correspond to patch I. Patch II has 19 residues; 7 overlap with the ClusPro residues; the last 7 residues are from patch II.

The third row provides interface residue predictions for GFD NAPP that were inferred from the fact that (1) the anti-GFD NAPP antibody 22C11 has a known GFD NAPP binding epitope (displayed) and (2) that 22C11 blocks the interaction between DR6 and GFD NAPP. Thus, we assume or predict that to block the DR6 interaction 22C11 is binding to the very GFD NAPP epitope that, at least in part, mediates binding to DR6.

Table 3: Comparisons between two different methods for predicting interface residues for the homology model of the DR6 ectodomain. More details are provided in the text and in Table 1. The good agreement exhibited between the two interface residue data sets tends to verify our DR6-GFD NAPP docked interaction model.

ClusPro predicted interface residues for DR6 ectodomain homology model:

F96,R98,H99,I103,E104,H107,D108,K120,L121,D128,E130,C131,T132,Q139,N141,A142,K158,E162, T163,E164,D165,R167

PPI-Pred predicted interface residues for DR6 ectodomain homology model

D108,K120,D128,E130,C131,T132,Q139

PPI-Pred produced three patch predictions for DR6. Only PPI-Pred predictions that agree with the ClusPro residues are underlined. For DR6, PPI-Pred produced three patch predictions (I, II, and III). Patch I is 27 residues, Patch II is 31 residues and patch III is 17 residues. All of the above displayed PPI-Pred residue predictions are derived from Patch II.

Methods

Recently, Nikolaev and co-workers described a protein-protein interaction between death cell receptor 6 (DR6), a transmembrane protein that is a member of the Tumor Necrosis Factor (TNF) family of proteins, and the N-terminal domain (1-286) of amyloid precursor protein (NAPP). The DR6-NAPP interaction is of considerable interest as it might play a role in Alzheimer's disease (AD) pathogenesis. Biophysical evidence suggests that the DR6-NAPP interaction specifically involves the DR6 ectodomain (67-211) and the growth factor-like domain (GFD) of NAPP. Clearly, a structural model of the DR6-GFD NAPP interaction would prove beneficial. We propose to build such a model through the integrated use of homology modeling, rigid-body protein-protein docking, energy minimization and computational binding affinity scoring and ranking. In particular, we constructed and optimized a homology model of the DR6 ectodomain, docked it to the crystal structure of GFD NAPP (28-123), and used a recently described empirical free energy function to select the best DR6-GFD NAPP protein-protein interaction model. During model construction, a number of independent methods were used to assess model quality and guide model selection. The final docked DR6-GFD NAPP model was subject to a number of tests against several lines of biophysical and theoretical data.

Building the DR6-GFD NAPP interaction model: computational details

Briefly, our modeling work-flow combined crystal structure evaluation (GFD NAPP), homology modeling (DR6), and homology model refinement through energy minimization. In an effort to ensure model structural quality, extensive quality tests were performed along the way. Ultimately, we relied on rigid-body protein-protein docking, rigid-body energy minimization and computational free energy scoring, to suggest a viable DR6-GFD NAPP interaction model for further testing and analysis.

Modeling and evaluating the structure of GFD NAPP. Rossjohn et al. have solved the crystal structure of residues 28-123 of GFD NAPP at 1.8 Å resolution. The structure is available in the Protein Data Bank (PDB) (http://www.pdb.org/pdb/home/home.do) with PDB identifier 1mwp. The high quality of the GFD NAPP crystal structure was verified using standard tools. We used the 1mwp structure in the present study and docked it to a homology model of the DR6

ectodomain[15]. All protein models and pictures were prepared using Swiss PDB Viewer (http://spdbv.vital-it.ch/)[34].

Modeling the structure of the ectodomain of DR6. We used the I-TASSER homology modeling server (http://zhang.bioinformatics.ku.edu/I-TASSER/) to construct a theoretical model of the DR6 ectodomain[16-18,22]. Towards that end, we submitted DR6 residues 67-211 to the I-TASSER server. The I-TASSER server builds homology models through an exhaustive process that involves automatic template selection, fragment reassembly of aligned regions, *ab initio* modeling of unaligned regions, clustering, energy evaluation and the optimization of a models hydrogen bonding network. Ultimately, we selected the top ranked I-TASSER DR6 model, based on the template supplied by the crystal coordinates of the neurotrophin receptor p75 in complex with the neurotrophin ligand (PDB code: 1sg1, chain X), for further analysis and eventual docking.

Evaluating the quality of the I-TASSER DR6 ectodomain homology model. Homology model construction is typically followed by visual and quantitative model evaluation. Importantly, the I-TASSER server automatically calculates and outputs various quality scores to assist end-users in model evaluation and selection. In particular, I-TASSER calculates an overall target quality score and a predicted target TM score and RMSD score. The quality of our DR6 ectodomain homology model was further assessed according to a Ramachandran map analysis and through the use of three independent server-based methods: ProSA, Qmean and DFIRE[34-38]. The ProSA server is available at: https://prosa.services.came.sbg.ac.at/prosa.php. Qmean and DFIRE were accessed through the SWISS-MODEL server (http://swissmodel.expasy.org/). All three servers use disparate methods to calculate quantitative scores that can be used to asses model quality and guide model selection.

Refining the structure of the I-TASSER ectodomain DR6 model. We used energy minimization, along with the Amber99 force field and the GB/SA implicit solvent model, to refine our DR6 ectodomain model. A termination criterion of 0.5 kcal/mol was applied and convergence was achieved. All calculations were carried out using the TINKER molecular modeling package

(http://dasher.wustl.edu/tinker/)[39]. The energy minimized I-TASSER DR6 ectodomain homology model was then used in the docking study.

Using docking to model the interaction between the ectodomain of DR6 and GFD NAPP. Our primary goal was to generate a reasonably good model of the interaction between the DR6 ectodomain and GFD NAPP. To achieve this goal we used the refined I-TASSER DR6 ectodomain model along with the GFD NAPP crystal structure as inputs to the ClusPro Docking server, version 1.0 (http://nrc.bu.edu/cluster/cluspro_v1.cgi)[23,24]. By default, the ClusPro server docks receptor (DR6) and ligand (GFD NAPP) structures using version 1.0 of the DOT rigid-body docking algorithm (http://www.sdsc.edu/CCMS/DOT/). The top 20,000 complexes generated by DOT are then filtered according to electrostatic and desolvation energies and the top 2,000 complexes are retained for further processing. The retained 2,000 conformations are then clustered according to interface RMSD and the top 10 docked models, following a short Charrm19 energy minimization, are made available for download. The top 10 ClusPro models probably capture most of the important rigid-body binding geometries and provide excellent starting structures for further refinement and analysis.

Refining, scoring and ranking the docked DR6-GFD models. Ultimately, we sought to narrow the top 10 ClusPro models down to a single physically realistic docked configuration. To accomplish this, the binding affinities of the top 10 ClusPro conformations were estimated in a hierarchical fashion. First, all 10 complexes were relaxed through rigid-body energy minimization using the Charmm19 force field. Next, a pseudo-binding affinity was calculated for all 10 models using the Charmm19 molecular-mechanics force field and GB/SA implicit solvent model (MM-GB/SA). All calculations were made using TINKER and default settings. Finally, ClusPro generated complexes with negative pseudo-binding affinities were scored using a recently described empirical free energy function that is available through CMDBioscience (http://www.cmdbioscience.com/)[25-27].

The CMDBioscience approach to protein-protein and protein-peptide binding free energy prediction (ΔG_{bind}) involves the use of a novel, fast, physics-based, empirical free energy function. The function is a six-term, regression-weighted expression and is given by,

$$\Delta G_{bind} = -0.79\Delta X_{+/-} + 0.075\Delta X_{c/s} - 0.65 X_{sb} - 0.86 X_{hb} - 0.00089 X_{gap} - 0.089\Delta X_{tor} - 0.33 \quad (1)$$

The first two terms refer to binding induced changes in the total number of solvent exposed charged (N-terminal nitrogen atoms, Arg and Lys side chain nitrogen atoms; O-terminal oxygen atoms, Asp and Glu side-chain carboxyl oxygen atoms; by default, His is treated as uncharged) and hydrophobic atoms (C and S atoms), respectively. The third and fourth terms refer to the total number of hydrogen bonds and the net number (difference between favorable and unfavorable charge-charge contacts) of short-range (≤ 4 Å) charge-charge or salt bridge interactions across the protein-protein interface. The contributions of these pairwise interface hydrogen bonding and salt bridge interactions are penalized according to the degree of solvent exposure, such that if the average solvent exposure is greater than some experimentally derived threshold value, energetic penalties are added to Eq. (1). The final three descriptors, in order, refer to the interface gap or void volume, the change in the number of solvent exposed side-chain torsions or the total number of side chain torsions buried at the interface, and a constant contribution. Changes in the number of solvent exposed main-chain torsions can also be counted for peptide ligands in the case of protein-peptide binding. Default values were used for each descriptor and all other important quantities. The model complex with the lowest empirical free energy score (predicted binding affinity) was ultimately selected as the best DR6-GFP NAPP structural interaction model.

Testing the DR6-GFD NAPP interaction model

The above described computational modeling work flow allowed us to identify a single, physically plausible, DR6-GFD NAPP interaction model or predicted complex structure. Importantly, our modeling work flow incorporated extensive *a priori* testing and analysis to ensure the fundamental physical reasonableness of our docked model. Our next task was to test and evaluate, *a posteriori*, the accuracy of the docked model against the available data. Towards that end, we tested our DR6-GFD NAPP model against a variety of biophysical data sets and a novel theoretical data set.

Testing the DR6-GFD NAPP interaction model against the available biophysical data. Initial biophysical model testing involved a quantitative comparison between the theoretical binding affinity calculated from our docked model and an experimentally determined value for the binding affinity. Biophysical model testing also involved the analysis of pairwise sequence alignments between GFD NAPP and the N-terminus of APLP2 (Amyloid beta (A4) precursor-like protein 2), on the one hand, and DR6 and p75, on the other. The alignments were judged to be worth analyzing because Nikolaev and co-workers also showed that an N-terminal fragment of APLP2 binds DR6 and p75 binds NAPP. Finally, in the case of GFD NAPP a comparison was made between the GFD NAPP interface residues implied by the docked model with the GFD NAPP interface residues suggested by the anti-GFD NAPP 22C11 antibody data.

Testing the DR6-GFD NAPP interaction model against the available theoretical data. The theoretical testing involved a comparison between the interface residues derived from our docked model with predicted binding site or interface residues for GFD NAPP and DR6, respectively, which were calculated using the protein-protein interaction prediction server (PPI-Pred) (http://bmbpcu36.leeds.ac.uk/ppi_pred/)[40]. From the coordinates of a monomeric protein structure, PPI-Pred typically predicts two or three binding patches or two or three well-defined residue patches that serve as protein-protein interaction sites. As will be discussed below, in the case of GFD NAPP PPI-Pred produced two patch predictions (I and II); in the case of the DR6 ectodomain, three predicted interface patches were generated (I, II and III).

References

1. Association, A.s. 2009 Alzheimer's disease facts and figures. *Alzheimers Dement* 5, 234-70 (2009).
2. Calissano, P., Matrone, C. & Amadoro, G. Apoptosis and in vitro Alzheimer disease neuronal models. *Commun Integr Biol* 2, 163-9 (2009).
3. Schindowski, K., Belarbi, K. & Buee, L. Neurotrophic factors in Alzheimer's disease: role of axonal transport. *Genes Brain Behav* 7 Suppl 1, 43-56 (2008).
4. Niewiadomska, G., Mietelska-Porowska, A. & Mazurkiewicz, M. The cholinergic system, nerve growth factor and the cytoskeleton. *Behav Brain Res*.
5. Alberghina, L. & Colangelo, A.M. The modular systems biology approach to investigate the control of apoptosis in Alzheimer's disease neurodegeneration. *BMC Neurosci* 7 Suppl 1, S2 (2006).
6. Guo, H. et al. Active caspase-6 and caspase-6-cleaved tau in neuropil threads, neuritic plaques, and neurofibrillary tangles of Alzheimer's disease. *Am J Pathol* 165, 523-31 (2004).
7. Guo, Q. et al. Alzheimer's presenilin mutation sensitizes neural cells to apoptosis induced by trophic factor withdrawal and amyloid beta-peptide: involvement of calcium and oxyradicals. *J Neurosci* 17, 4212-22 (1997).
8. LeBlanc, A.C. The role of apoptotic pathways in Alzheimer's disease neurodegeneration and cell death. *Curr Alzheimer Res* 2, 389-402 (2005).
9. Capsoni, S., Giannotta, S. & Cattaneo, A. Nerve growth factor and galantamine ameliorate early signs of neurodegeneration in anti-nerve growth factor mice. *Proc Natl Acad Sci U S A* 99, 12432-7 (2002).
10. Hock, C., Heese, K., Hulette, C., Rosenberg, C. & Otten, U. Region-specific neurotrophin imbalances in Alzheimer disease: decreased levels of brain-derived neurotrophic factor and increased levels of nerve growth factor in hippocampus and cortical areas. *Arch Neurol* 57, 846-51 (2000).
11. Nagahara, A.H. et al. Long-term reversal of cholinergic neuronal decline in aged non-human primates by lentiviral NGF gene delivery. *Exp Neurol* 215, 153-9 (2009).
12. Matrone, C., Ciotti, M.T., Mercanti, D., Marolda, R. & Calissano, P. NGF and BDNF signaling control amyloidogenic route and Abeta production in hippocampal neurons. *Proc Natl Acad Sci U S A* 105, 13139-44 (2008).
13. Nikolaev, A., McLaughlin, T., O'Leary, D.D. & Tessier-Lavigne, M. APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. *Nature* 457, 981-9 (2009).
14. Berg, T. Small-molecule inhibitors of protein-protein interactions. *Curr Opin Drug Discov Devel* 11, 666-74 (2008).
15. Rossjohn, J. et al. Crystal structure of the N-terminal, growth factor-like domain of Alzheimer amyloid precursor protein. *Nat Struct Biol* 6, 327-31 (1999).
16. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5, 725-38.
17. Zhang, Y. I-TASSER: fully automated protein structure prediction in CASP8. *Proteins* 77 Suppl 9, 100-13 (2009).
18. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40 (2008).
19. Chopra, G., Summa, C.M. & Levitt, M. Solvent dramatically affects protein structure refinement. *Proc Natl Acad Sci U S A* 105, 20239-44 (2008).
20. Summa, C.M. & Levitt, M. Near-native structure refinement using in vacuo energy minimization. *Proc Natl Acad Sci U S A* 104, 3177-82 (2007).
21. Summa, C.M., Levitt, M. & Degrado, W.F. An atomic environment potential for use in protein structure prediction. *J Mol Biol* 352, 986-1001 (2005).

22. Zhang, Y. Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins* 69 Suppl 8, 108-17 (2007).
23. Comeau, S.R., Gatchell, D.W., Vajda, S. & Camacho, C.J. ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic Acids Res* 32, W96-9 (2004).
24. Comeau, S.R., Gatchell, D.W., Vajda, S. & Camacho, C.J. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 20, 45-50 (2004).
25. Audie, J. Continued development of an empirical function for predicting and rationalizing protein-protein binding affinities. *Biophys Chem* 143, 139-44 (2009).
26. Audie, J. Development and validation of an empirical free energy function for calculating protein-protein binding free energy surfaces. *Biophys Chem* 139, 84-91 (2009).
27. Audie, J. & Scarlata, S. A novel empirical free energy function that explains and predicts protein-protein binding affinities. *Biophys Chem* 129, 198-211 (2007).
28. Comeau, S.R. et al. ClusPro: performance in CAPRI rounds 6-11 and the new server. *Proteins* 69, 781-5 (2007).
29. Comeau, S.R., Vajda, S. & Camacho, C.J. Performance of the first protein docking server ClusPro in CAPRI rounds 3-5. *Proteins* 60, 239-44 (2005).
30. Duan, Y., Reddy, B.V. & Kaznessis, Y.N. Physicochemical and residue conservation calculations to improve the ranking of protein-protein docking solutions. *Protein Sci* 14, 316-28 (2005).
31. Li, L., Chen, R. & Weng, Z. RDOCK: refinement of rigid-body protein docking predictions. *Proteins* 53, 693-707 (2003).
32. Bradford, J.R. & Westhead, D.R. Asymmetric mutation rates at enzyme-inhibitor interfaces: implications for the protein-protein docking problem. *Protein Sci* 12, 2099-103 (2003).
33. Yi, H., Qiu, S., Cao, Z., Wu, Y. & Li, W. Molecular basis of inhibitory peptide maurotoxin recognizing Kv1.2 channel explored by ZDOCK and molecular dynamic simulations. *Proteins* 70, 844-54 (2008).
34. Guex, N. & Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714-23 (1997).
35. Benkert, P., Kunzli, M. & Schwede, T. QMEAN server for protein model quality estimation. *Nucleic Acids Res* 37, W510-4 (2009).
36. Benkert, P., Tosatto, S.C. & Schomburg, D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* 71, 261-77 (2008).
37. Wiederstein, M. & Sippl, M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35, W407-10 (2007).
38. Zhou, H. & Zhou, Y. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Sci* 11, 2714-26 (2002).
39. Ponder, J.W. et al. Current status of the AMOEBA polarizable force field. *J Phys Chem B* 114, 2549-64.
40. Bradford, J.R. & Westhead, D.R. Improved prediction of protein-protein binding sites using a support vector machines approach. *Bioinformatics* 21, 1487-94 (2005).