Chapter VI: The Structure and Function of the Human G2A Receptor: Collaboration Between Theory and Experiment

Chapter VI: The Structure and Function of the Human G2A Receptor: Collaboration Between Theory and Experiment

Abstract:

Lysophospholipids (LP) regulate a wide range of cellular responses including proliferation, apoptosis, cell motility and migration [Fukushima 2001]. These molecules have long been involved in the pathogenesis of inflammatory, autoimmune and neoplastic diseases [Huang 2002] but until recently they have not been linked to specific cell-surface receptors. The discovery, in the late 90's, of the first LP receptor gene encoding a GPCR (Hecht 1996) has given a considerable boost to research in the field. The impetus is mostly provided by the widely accepted idea that understanding these lipid mediators and their receptors may lead to the development of novel therapeutic approaches [Brinkmann 2002]. Herein we present results of a recent joint theory/experimental study of the structure and function of G2A, an immunoregulatory GPCR specific for the proinflammatory lipid lysophosphatidylcholine (LPC) [Chisolm 1996].

Introduction:

1. G2A, an immunoregulatory GPCR with lysophosholipid specificity

1.1 Discovery of G2A and signal transduction

G2A is a GPCR cloned in Owen Witte's lab at UCLA as a transcriptional target of the leukemogenic tyrosine kinase BCR-ABL [Weng 1998]. Depending on the cell type, upon overexpression, G2A was shown to induce pleiotropic effects on cell cycle, survival and cytoskeleton dynamics (**Table 6-1**).

Table 6-1. Multiple cellular effects induced by G2A overexpression

Cell type	Effect	Reference
NIH 3T3 fibroblasts	G2/M accumulation and block in the progression of mitosis	Weng, Z. et al, PNAS, 1998
Swiss 3T3 fibroblasts	Actin reorganization into stress fibers mediated by $G\alpha_{13}$ and RhoA	Kabarowski JHS et al, PNAS, 2000
NIH 3T3	Loss of contact inhibition and anchorage-independent survival	Zohn, IE et al, Oncogene, 2000
HeLa	Apoptosis via $G\alpha_{13}$ and $G\alpha_s$ mediated pathways	Lin P, Ye RD, JBC, 2003

1.2 G2A deficient mice develop spontaneous autoimmunity

G2A deficient mice were generated in Owen Witte's lab by conventional gene targeting technology [Le 2001]. Hematopoietic cells from these mice were found to have an increased

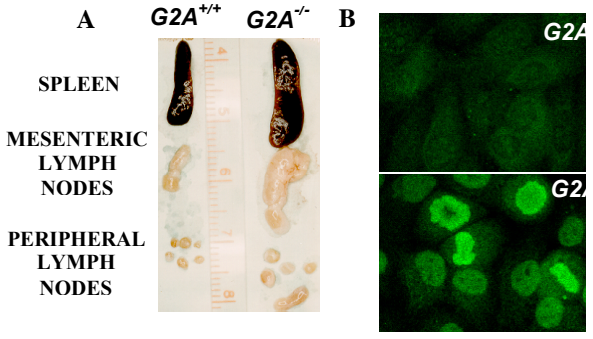

Fig 1. Autoimmunity in *G2A*$^{-/-}$ mice (A) Enlargement of lymphoid organs in *G2A*$^{-/-}$ mice; **(B** presence in G2A$^{-/-}$ mice of serum autoantibodies reacting against nuclear antigens

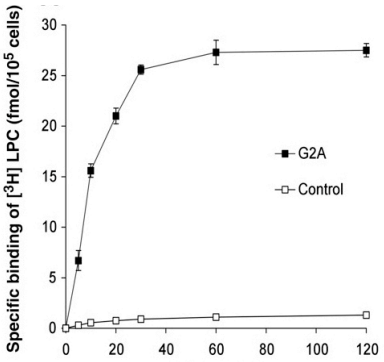

Figure 6-2. Time dependence of [^{3}H] LPC binding to cell homogenates from HEK 293 EGFP (control) or HEK 293 G2A-EGFP (G2A) cells

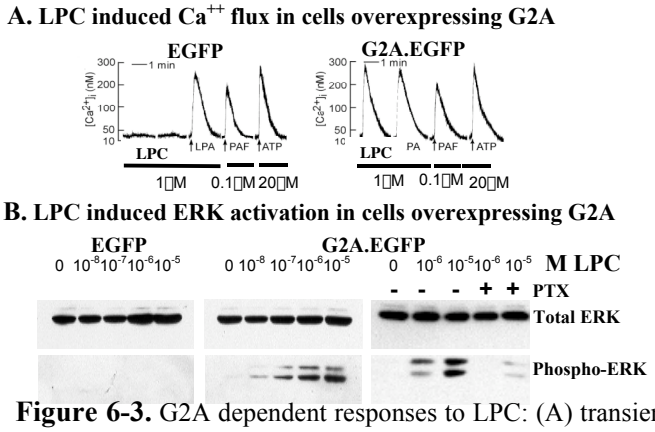

Figure 6-3. G2A dependent responses to LPC: (A) transient increases in intracellular calcium concentration in G2A-expressing MCF10A cells; (B) Gα_i dependent activation of ERK MAP kinase in G2A-expressing CHO cells

susceptibility to malignant transformation by BCR-ABL [Le 2002]. However, young *G2A*$^{-/-}$ mice appear normal and exhibit no discernible histological abnormalities of their hematopoietic and lymphoid tissues. As they age (>1.5 yrs), *G2A*$^{-/-}$ mice spontaneously develop an autoimmune syndrome characterized by progressive enlargement of secondary lymphoid organs (**Figure 6-1A**), lymphocytic infiltration in the lungs and liver, increased IgG levels, deposition of immune complexes in glomeruli and high levels of serum anti-nuclear antibodies (Fig 1B) [Le 2001]. These features are reminiscent of the human disease, systemic lupus erythematosus (SLE) [Nishimura 1999]. The only immunological abnormality found so far in young *G2A*$^{-/-}$ mice that could potentially explain the autoimmune syndrome is represented by increased proliferation and sensitivity of T lymphocytes from these mice following activation [Le 2001 and C.G.Rado and O.N.Witte, unpublished].

1.3 Lysophosphatidylcholine is a ligand for G2A

The serum-borne bioactive lysophospholipid lysophosphatidylcholine (LPC) was identified as a ligand for G2A (Kabarowski 2001) (**Figure 6-2**) and shown to elicit intracellular calcium release and ERK MAP kinase activation via Ga_i heterotrimeric G proteins (**Figure 6-3**).

LPC is produced from low-density lipoproteins (LDLs) and cell membrane derived phosphatidylcholine (PC) as a result of hydrolysis by phospholipase A_2 (PLA_2) (McMurray 1993). As a component of oxidized low-density lipoprotein (oxLDL) LPC plays an important etiological role in atherosclerosis [Lusis 2000] and is implicated in the pathogenesis of SLE [Koh 2000]. While aged $G2A^{-/-}$ mice develop a lupus-like disease [Le 2001], the role of LPC in this process is currently unknown.

1.4 G2A regulates migration of T lymphoid cells to LPC

Several lines of evidence suggest a role for the LPC-G2A ligand-receptor pair in regulating chemotaxis: microinjection studies in Swiss 3T3 cells demonstrate that G2A can couple to cytoskeletal effectors such as RhoA [Kabarowski 2000] and overexpression of G2A in the human T cell line Jurkat enables these cells to migrate towards LPC [Kabarowski 2001]. LPC has also been shown to be a chemotactic factor for primary monocytes and T lymphocytes but the receptor involved has not been identified [McMurray 1993, Quinn 1988].

To unequivocally demonstrate that G2A can mediate the chemotactic effect of LPC, expression of this GPCR was chronically suppressed in the T lymphoid cell line DO11.10 [White 1983]. This was accomplished by retroviral transduction of RNAi with co-linked fluorescent markers (**Figure 6-4**) [C.G.Rado and O.N.Witte, unpublished].

G2A suppression significantly impaired chemotaxis to LPC (**Figure 6-5A**). It did not affect the response to SDF1-a, which is mediated through the chemokine receptor CXCR4 [Berger 1999] (**Figure 6-5B)**. Chemotaxis of WT DO11.10 cells to LPC can be further enhanced by retroviral transduction of G2A. The amount of G2A in retrovirally transduced cells exceeds the endogenous level by approximately 20 fold (**Figures 6-5 C, D**).

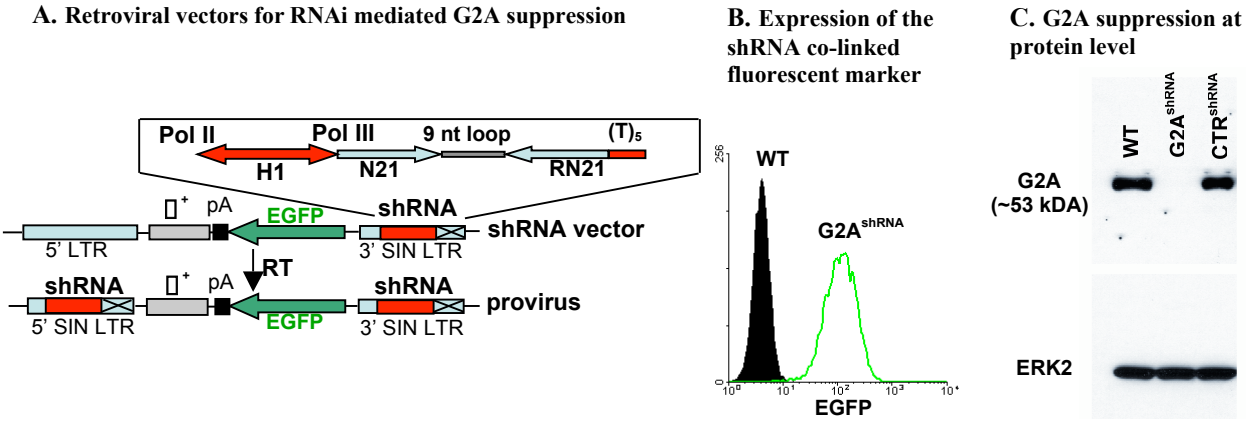

Fig. 4 Silencing of G2A in DO11.10 cells. (A) The bi-directional human H1-RNA promoter (H1) coordinates expression of the short hairpin RNA (shRNA, RNA pol III dependent) and of EGFP (RNA pol II dependent). Reverse transcription (RT) results in the duplication of the shRNA cassette inserted in the 3' self inactivating LTR (3' SIN LTR); $(T)_5$ – termination signal for the RNA pol III; pA-polyadenylation signal. **(B)** Expression of the EGFP co-linked marker by retrovirally transduced DO11.10 cells. **(C)** Expression levels of G2A in DO11.10 T cells transduced with G2A specific ($G2A^{shRNA}$) or control (CTR^{shRNA}) encoding retroviruses (Western Blot using the rabbit polyclonal antibodies against G2A; ERK2 blot indicates equivalent total protein amounts).

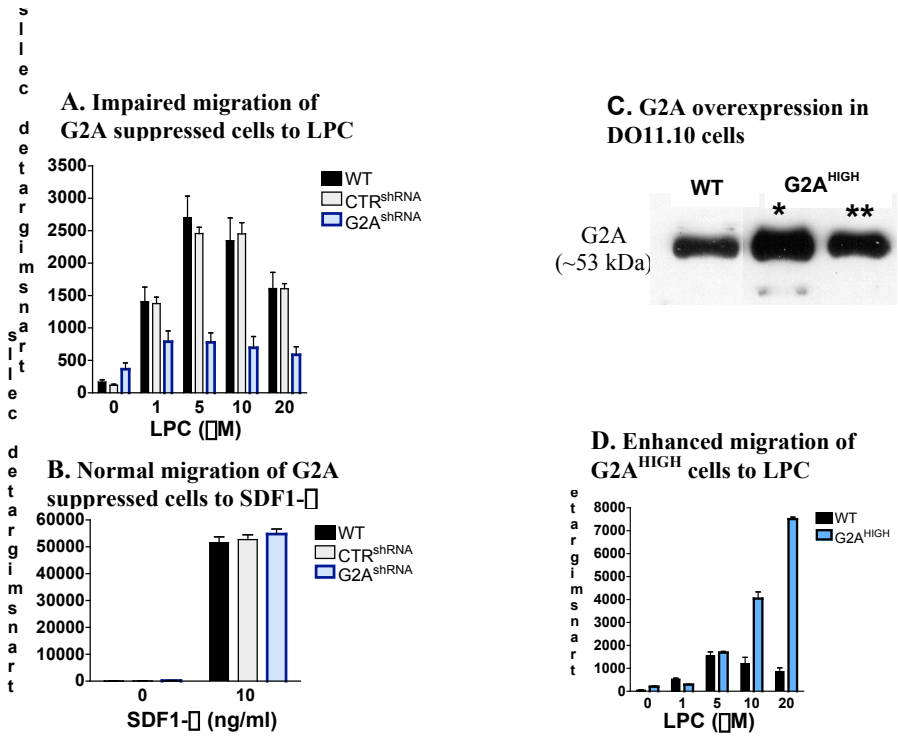

Fig. 5 LPC is a chemotactic factor for DO11.10 cells and this effect is dependent on G2A levels. $2x10^5$ WT and $G2A^{shRNA}$ or control (CTR^{shRNA}, corresponding to a target sequence specific for human TDAG8) cells were washed 3 times with serum-free medium containing 0.1% fatty acid free BSA, mixed and added to the upper chamber of a 24 well plate with 5.0 μm pore size polycarbonate membranes (Costar); LPC **(A)** or SDF1-α **(B)** were added to the lower chamber and the plate was incubated for 2 hr at 37ºC in a 8% CO_2 incubator; **(C)** Western Blot to estimate the amount of G2A in cells overexpressing the receptor ($G2A^{HIGH}$). Lysates from $G2A^{HIGH}$ cells were diluted 10 (*) and 20 (**) fold before loading on SDS-PAGE; **(D)** Transmigration of WT and $G2A^{HIGH}$ cells to LPC.

1.5 A potential role for G2A in chemotaxis of macrophages towards apoptotic cells

During apoptosis, caspase-3 mediated activation of the calcium-independent cytosolic phospholipase A_2 ($iPLA_2$) leads to production of LPC [Kim 2002, Lauber 2003]. This lysophospholipid can then act as a chemoattractant for macrophages, which are the cells responsible for efficient removal of the apoptotic bodies. If G2A would play a role in chemotaxis of macrophages to LPC, it is conceivable that clearance of LPC-releasing apoptotic cells is impaired in G2A deficient mice. In turn, this could lead to postapoptotic

necrosis, aberrant presentation of self antigens [Lauber 2003] and eventually systemic lupus-like autoimmunity. This hypothesis is currently being tested using $G2A^{-/-}$ mice. Significantly, overexpression of G2A in the monocytic/macrophage cell lines J774A.1 and U937 [Ralph 1976, Sundstrom 1976, L.Yi, C.G.Rado and O.N.Witte unpublished observations] renders these cells responsive to LPC induced chemotaxis. We plan to use this property to test the functional consequences of the predicted mutations in the G2A LPC binding site.

1.6 TDAG8 is a GPCR related to G2A by sequence homology and pattern of expression.

TDAG8 (T cell death-associated gene 8) was discovered by differential mRNA display as a gene strongly upregulated during glucocorticoid-induced apoptosis of thymocytes [Choi 1996]. Our interest in TDAG8 is motivated by its high degree of sequence homology with G2A (over 55% sequence similarity without carboxy and amino terminus) and by studies demonstrating that G2A and TDAG8 are co-expressed in lymphocytes and macrophages [C.G.Rado and O.N.Witte, unpublished observations]. Taken together, these findings suggest a possible functional connection between G2A and TDAG8.

The glycosphingolipid psychosine was recently proposed to activate TDAG8 leading to a block in cytokinesis and formation of giant multinucleated cells [Im 2001]. These effects are the hallmark of Globoid Cell Leukodystrophy (GLD) an autosomal recessive sphingolipidosis caused by deficient activity of the lysosomal hydrolase galactosylceramide beta-galactosidase (GALC) leading to accumulation of psychosine and widespread destruction of oligodendroglia in the CNS and to subsequent

demyelination [Im 2001]. However, given the non-physiological concentrations of lipid required to activate TDAG8, it is still unclear if psychosine does indeed represent the natural ligand for this GPCR [Mitchison 2001]. TDAG8 is therefore a GPCR that, in contrast to G2A, is less well characterized in terms of function and ligand specificity. It is of interest to us to determine if structural modeling of TDAG8 based on the 3D structure of G2A will allow "virtual screening" for novel ligands for this enigmatic GPCR.

2. Computational methods for predicting the structure and function of GPCRs

2.1 Overview of MembStruk and HierDock ab initio methods and comparison with other GPCR modeling methods-

2.1 Homology methods for modeling GPCRs:

The difficulty in generating 3D structures for GPCRs is in obtaining high quality crystals of these membrane-bound proteins for high resolution X-ray diffraction data. It is equally difficult to use NMR to determine 3D structures of GPCRs. Hence it is widely accepted that theory and computation to predict the 3D structures of GPCRs from first principles, can aid the structure based drug design for many GPCR targets [for example Strader 1994, Parrill 2000 and many other references for different GPCRs]. Successful protein structure prediction methods for globular proteins generally utilize homology to known structures [John 2003]. This is not practical for GPCRs (with just one crystal structure). Moreover homology derived models are not reliable when the sequence homology is very low below 30% or less (in the "twilight zone") [Rost 1999, Chung 1996, Brenner 1998]. The sequence identity between G2A human and bovine rhodopsin

is 13% over the whole sequence, and hence in the "twilight zone" where homology methods are known to fail.

GPCR structures have also been modeled using the properties of conserved residues in multiple sequence alignments followed by optimization of the structure using distance restraint to maximize the hydrogen bonds [Lomize 1999]. Shacham et al have also predicted the structure of bovine rhodopsin using an approach based on specificity of protein-protein interaction and protein-membrane interaction and the amphiphathic nature of the helices. However there is not much detail of their method available in literature [Shacham 2001].

2.2 MembStruk – an ab initio GPCR structure prediction method:

GPCRs have a well defined three dimensional topology, with seven helical transmembrane (TM) domains, and this could be an advantage for first principles methods because it provides an organizing principle allowing some of the structural information to be deduced from sequence. MembStruk method [Vaidehi 2002, and in Trabanino 2003 submitted], is an *ab initio* structure prediction algorithm using no information from the high resolution crystal structure of rhodopsin or bacteriorhodopsin. A simple flow chart of the method is shown in **Figure 6-6.**

A detailed description of the MembStruk *ab initio* GPCR modeling methodology consists of the following steps [the details are in Vaidehi 2002 and Trabanino 2003 submitted]:

- TM Prediction: Predict the seven TM domains using hydropathicity analysis "TM2ndS" [Trabanino 2004] combined with information from sequence alignments. The extent of the TM regions are predicted using sequence alignments of sequences varying from 40% to over 90% sequence variability as input. The second step of TM2ndS is to calculate the consensus hydrophobicity for every residue position in the alignment using the average hydrophobicity of all the amino acids in that position over all the sequences in the multiple sequence alignment. Then, we calculate the average hydrophobicity over a *window size (WS)* of residues about every residue position, using WS ranging from 12 to 20. This average value of hydrophobicity at each sequence position is plotted to yield the hydrophobic profile, for WS=14. The baseline for this profile serves as the threshold value for determining the TM regions.
- Position of maximum hydrophobicity: Identify lipid-accessible residues from the sequence alignments (as the less conserved residues) and from analysis of the peaks in hydrophobicity [Jayasinghe 2001, Eisenberg 1984] of the hydrophobic residues in the sequence.
- Optimization of helical kinks: Construct canonical helices for the predicted TM segments and optimize the structures of the individual helices with energy minimization followed by fast torsional NEIMO dynamics [Jain 1993, Vaidehi 1996]. This optimizes the bends and kinks in each helix.
- Assemble the helix bundle: The helical axes are oriented according to the 7.5 Å electron density map of frog rhodopsin [Schertler, 1998], depending on sequence

similarity. The relative translational orientation of each helix is based upon forming the best fitting plane of all the hydrophobic centers obtained from step 2.

- Monte Carlo Optimization of rotation and translational degrees of freedom: This step is an important step that optimizes the rotational and translational degrees of freedom of each helix with respect to the other. Here optimize the rotational and translational orientation of the helices using a systematic search algorithm over a grid of rotational angles and translational distances. This step allows the system to surmount energy barriers. Coarse grain optimization of the helical orientations is performed using the net hydrophobic moment of the middle one-third of the helix about their hydrophobic centers.

- Optimization of the assembled helical bundle in explicit lipids: Embed the helix bundle into a lipid bilayer and optimize the composite system. Equilibration of the helix bundle plus lipid bilayer system uses Rigid Body Molecular Dynamics [Ding 1992, Lim 1997]. The helix bundle surrounded by lipid bilayers was optimized using rigid body dynamics with DREIDING forcefield [Mayo 1990] and CHARMM22 [Mackerrell 1998] charges for the protein.

- Optimization of the final model: Construction of the inter-helical loops and disulfide bridges using Whatif [Vriend 1990]. Optimization of the final model in presence of lipid bilayers.

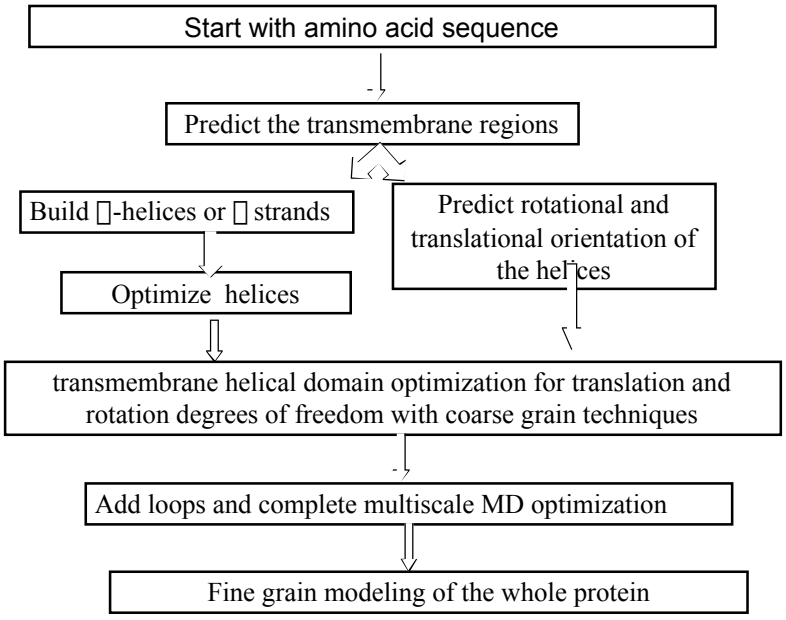

Fig. 6 : Flow Chart of MembStruk, the *ab initio* method for predicting GPCR structures.

2.3 Validation: Structure prediction for bovine rhodopsin: The only GPCR with an experimental 3D crystal structure is bovine rhodopsin [Palczewski 2000; Teller 2001]. Thus, this is the only structure with which to compare MembStruk predictions with experiment. The TM regions of the predicted structure for rhodopsin agree with the crystal structure, to 2.8Å CRMS for the main chain atoms [Trabanino 2004]. Comparing the individual helices lead to CRMS errors of 1.0Å for TM2, 2.1Å for TM2, 1.2Å for TM3, 1.1Å for TM4, 1.8Å for TM5, 2.2Å for TM6 and 1.6Å for TM7. This excellent agreement with the crystal structure indicates that the MembStruk procedure predicts the helical regions well, without using any knowledge of the crystal structure. These results have just been communicated to the Biophysical Journal [Trabanino 2004].

***2.4 Function Prediction Methods for GPCRs**:* The structures of GPCRs, derived using the methods described above, are used to predict the binding site and affinity for various ligands for which there is experimental data available to validate the structure predictions. This HierDock protocol has been applied successfully to dopamine binding to human D2 dopamine receptor [Kalani 2004], epinephrine to β2-adrenergic receptor [Freddolino 2004, odorants binding to mammalian olfactory receptors [Floriano 2000 and Floriano 2004, Hall 2004 submitted], amino acid discrimination in aminoacyl t-RNA synthetases [Wang 2002, Zhang 2002, Datta 2003 submitted, Kekenes-Huskey 2003] and other globular proteins [Floriano 2003, Datta 2002, Datta 2003].

The HierDock Protocol: The HierDock ligand screening protocol follows a hierarchical strategy for examining ligand binding conformations, and calculating their binding energies. The steps are as follows:

1. Coarse grain docking: First we carry out a coarse grain docking procedure to generate a set of conformations for ligand binding in the receptor. Currently we use DOCK 4.0 [Ewing 1997] to generate and score 1000 configurations, of which 10% (100) were selected for further analysis. We are developing a new approach here, MPSim-Dock [Wendel, Vaidehi and Goddard unpublished] that we believe will be faster and more accurate.

2. Ligand optimization: The 100 best conformations selected for each ligand from step 1 are subjected to all-atom minimization keeping the protein fixed but the ligand movable. The solvation of each of these 100 minimized structures is calculated using a Poisson-Boltzmann based continuum solvation method [Tannor 1994]. Currently

we use the Surface Generalized Born (SGB) continuum solvation method [Ghosh 1998] or the Analytical Volume Generalized Born (AVGB) method [Zamanakos 2001]. The binding energies (BE) are calculated using

$$BE = PE \text{ (ligand in solvent)} - PE \text{ (ligand in protein)} \quad (1)$$

Then the 10 structures based both on binding energies and buried surface areas are selected from these 100 structures for the next step.

3. Monte Carlo Optimization of the ligand bound conformations with flexible receptor binding site: In this step we use Monte Carlo method to generate various possible ligand conformations in the field of the protein. The conformations are selected based on diversity of the conformations from each other to cover the conformational space. We call these conformations as "family heads" and they differ from each other at least by 0.6Å in CRMS (RMSD in coordinates). Next the energy of interaction of each family head with the protein is calculated, and about 10% of the good energy "family heads" are selected for further enrichment of these conformations. The enrichment is done by generating conformations using Monte Carlo procedure and selecting only those conformations that are close (within 0.6Å CRMS) to the good energy family heads. A ligand conformation that is within 0.6Å of the family head is known as a child of that family head. The enrichment cycle is performed until each chosen good energy family head gets at least 6 children on an average. We then calculate the ligand protein interaction energy for all the children of each family head. The conformations (family heads and children) are all sorted by energy and the best 10 ligand conformations are chosen for the next step of protein movable optimization. The binding energy calculated with these conformations show good agreement with

the measured binding affinities [Kekenes-Huskey, Vaidehi and Goddard in preparation].

4. Side chain optimizations of the residues in the binding site**:** For each ligand conformation chosen in the previous step, we map out all the residues within 5Å of the binding site. The side chain rotamer conformations of each one of these mapped residues (within 5Å) are placed optimally in response to the ligand protein interaction energy using the side chain placement method called SCREAM being developed in the Goddard laboratory. SCREAM uses a side-chain rotamer library (1478 rotamers with 1.0Å resolution) and uses the all-atom DREIDING energy function with AVGB continuum solvation method to evaluate the energy of interaction of each side chain rotamer with the ligand and the rest of the protein. This gives excellent results for side chain placement for many test cases [Kam, Vaidehi and Goddard - publication in preparation]. Once the side chains of all the residues are optimized, the potential energy of the receptor/ligand complex, is minimized using conjugate gradient minimization technique to a convergence of 0.1kcal/mol/Å in force for an atom. Subsequently the binding energy is calculated using equation (1) and the 5 best structures of the receptor/ligand complex structures are examined for good salt bridges, hydrogen bonds and other hydrophobic contacts. The energy contribution from each residue in the binding pocket, to ligand binding is calculated. Next we optimize the structure of the receptor/ligand complex, allowing the structure of the protein to accommodate the ligand. This all-atom receptor/ligand energy minimization is essential to identify the optimum conformations for the complex, and it is performed on the 10 structures from the previous step. Using these optimized

structures, we calculate the binding energy as the difference between the energy of the ligand in the protein and the energy of the ligand in water. The energy of the ligand in water is calculated using DREIDING FF and the SGB or AVGB continuum solvation method.

2.5 Scanning the entire receptor for the binding region: The above HierDock procedure is efficient for sampling a region of volume ~ 1000 to 2000 A^3 (a cube with sides of 10 to 14Å). However, for GPCRs the binding region is not known this well. Thus our first step on a new GPCR structure is ScanBindSite. In the procedure we scan the entire protein for potential binding regions with no assumption on the binding site. The entire molecular surface of the predicted structure is mapped [Connolly 1983] and spheres representing the empty volume of the receptor are generated (currently using the Sphgen program in DOCK4.0 suite of programs). The entire set of receptor spheres is partitioned into ~10 to 15 overlapping regions and a set of known agonists and antagonists are used to scan for the putative binding region. This scan uses only the first 2 steps of the HierDock protocol described above. The consensus of ligand structures corresponding to the most energetically favorable sites is used to determine the *putative binding region* with ~ 1000 to 2000 A^3. In some cases of long ligands, scanning the entire receptor for putative binding regions gave two or three possible binding regions with similar binding energies in each region. In such cases we merged the spheres of all the regions with similar binding regions and performed HierDock into this large region.

2.6. Determination of binding site and binding energy for all ligands:

After determining the putative binding region, we carry out a full HierDock analysis (steps 1 to 4 as described above) using this region to determine the binding site and binding energy of the list of agonists and antagonists. Sometimes the HierDock procedure for agonists could be performed in different region than antagonists. This depends on how the putative binding region was derived from the previous step. The resulting site is compared to any available mutation data to evaluate the predicted binding site. If there are problems in those residues that are known to directly recognize the ligand, then we may go back to the last step of the HierDock process to examine the next best binding energy structures. If there are still problems we may have to go back to previous HierDock levels to find good structures. So far this iteration of earlier structures has not been necessary.

2.7 Validation for function prediction protocol:

HierDock has been used to predict the binding site for aminoacyl t-RNA synthetases [Wang et al 2002, Kekenes-Huskey 2003] and 37 other co-crystals of globular proteins [Datta, 2002, Floriano 2003]. We also have used HierDock for preliminary study of several GPCRs [Floriano 2000; Vaidehi 2002]. We recently validated HierDock for binding of 11-cis-retinal to bovine rhodopsin [Trabanino 2004]. The CRMS between crystal structure and the docked structure for cis-retinal is 0.6Å, which is excellent considering that no information of the binding site was assumed. Using the predicted protein structure for rhodopsin (instead of the experimental structure)

we still find the binding site of 11cis-retinal and the CRMS to the crystal structure is 2.8Å in the docked structure.

3. Significance of the collaborative work: This proposed tight collaboration between experiment and theory on GPCRs such as G2A and TDAG8 would provide insights into the binding site of ligands for these receptors. The experimental work will be largely guided by the theory and hence will allow an enormous reduction in the number of experiments required to identify the active site. Each experiment will be targeted based on the predicted structure. The experimental results in turn would be used to refine the prediction methods in a generic fashion. Each group has long experience and expertise in their respective research areas that we believe that this collaboration would be extremely productive in reaching the goals stated in this proposal. The preliminary results of this collaboration on G2A receptors indicate sufficient promise to justify a focused effort. In addition to the importance of better understanding this complex G2A GPCR system, we believe that this work would illustrate how to couple the new computational methods effectively to experiment to determine the structure and function of other GPCRs.

C. Results and Discussion:

1. Structure and function prediction for G2A

The three-dimensional structure of human G2A structure was predicted using a previous version of MembStruk3.0 that did not have the optimization of the translational degree of freedom of the helices. The most current method is however described in the background section. Subsequently we used HierDock method to scan the entire receptor

and predict the binding site of LPC in G2A structure. The location of the binding site of LPC in G2A is shown in **Figure 6-7**. LPC binds to the human G2A with a binding affinity of 9.0 kcal/mol (where positive binding energy is better binding). The binding site of LPC is located between TM 3, 5, and 6 and includes residues from the EC2 loop.

2. Preliminary identification of key G2A residues contacting LPC:

The residues within 5.0Å of the bound structure of LPC is shown in **Figure 6-8.** There are three functional components to LPC ligand: the choline head, the phosphate group, hydroxy-modified middle region with the non-polar tail. The phosphate head is sandwiched between two charged residues while the hydrophobic tail is embedded in a channel for hydrophobic residues.

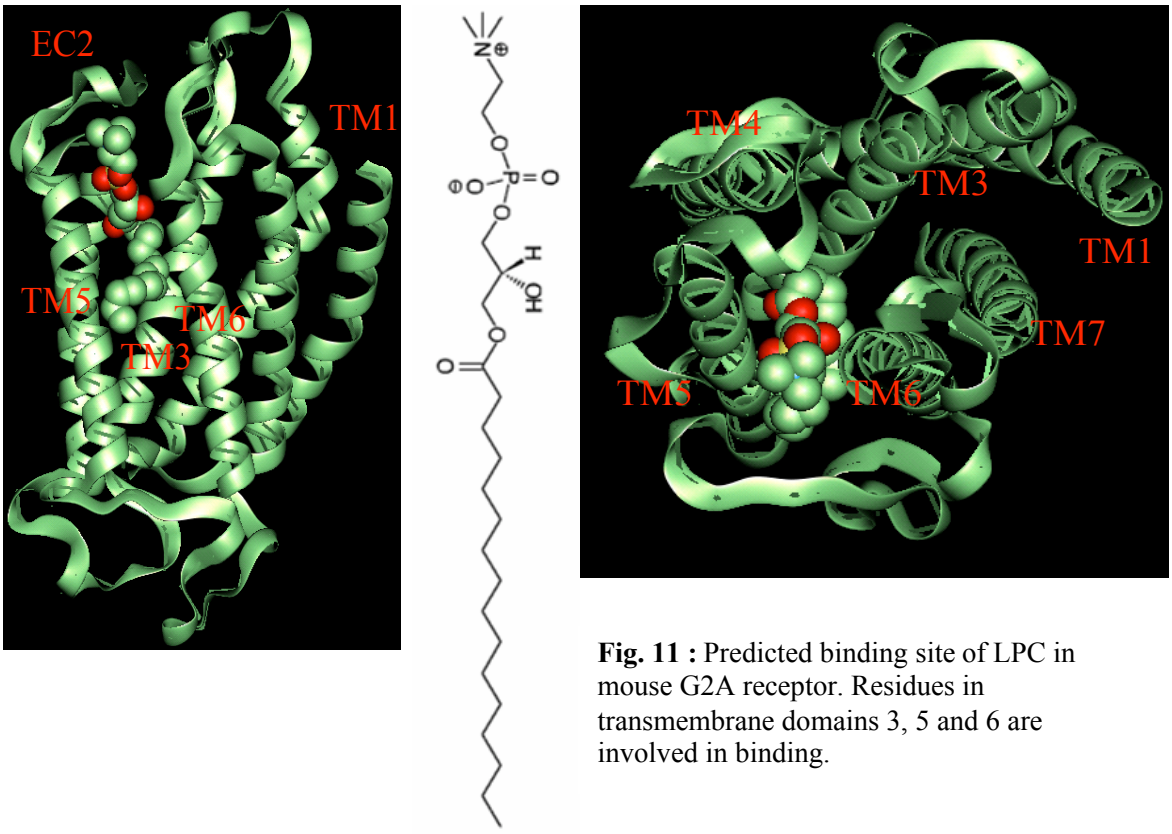

Fig. 11 : Predicted binding site of LPC in mouse G2A receptor. Residues in transmembrane domains 3, 5 and 6 are involved in binding.

2.1 Residues in contact with the choline moiety of LPC

The binding site of the choline moiety in the G2A receptor is shown in **Figure 6-8A**. The choline group clearly prefers the top residues in TM3, 5, and 6 and several residues in the EC2 as shown in Fig. 13A. The hydroxyl group of the choline group forms a hydrogen bond with Y199. The choline-binding site is predominantly hydrophobic and includes residues Y120, F178, Q179, F187, M189 and L190.

2.2 Residues in contact with the phosphate group of LPC

The phosphate group of LPC is negatively charged at the physiological pH of 7.4. The negatively charged phosphate is stabilized by two positively charged residues, Arg203 in TM5 and Lys265 in TM6 as shown in **Figure 6-8B**. Arg203 forms a tight salt bridge with a distance of 3.1 Å and Lys265 forms a 3.6 Å salt bridge to the phosphate.

2.3 Residues in contact with the alkyl chain of LPC.

The alkyl tail region of LPC is hydrophobic and hence stabilized by a stretch of hydrophobic amino acids in the TM barrel of the G2A receptor. The side chains of TM5 and 6 provide an aliphatic pocket for the alkyl tail as shown in **Figure 6-8C**. The residues in the hydrophobic channel are Val206 (TM5), Ile210 (TM5), Pro211 (TM5), Ile214 (TM5), Phe255 (TM6), Tyr258 (TM6), His259 (TM6), and Val262 (TM6). However we find that there are also some polar residues present in this hydrophobic channel in the vicinity of the alkyl tail of LPC. The presence of these polar residues hints at the possibility of some modified lipid binding to this receptor. The location of His259

especially causes us to speculate that the lipid could be modified to make use of His259 contact as hydrogen bond donor.

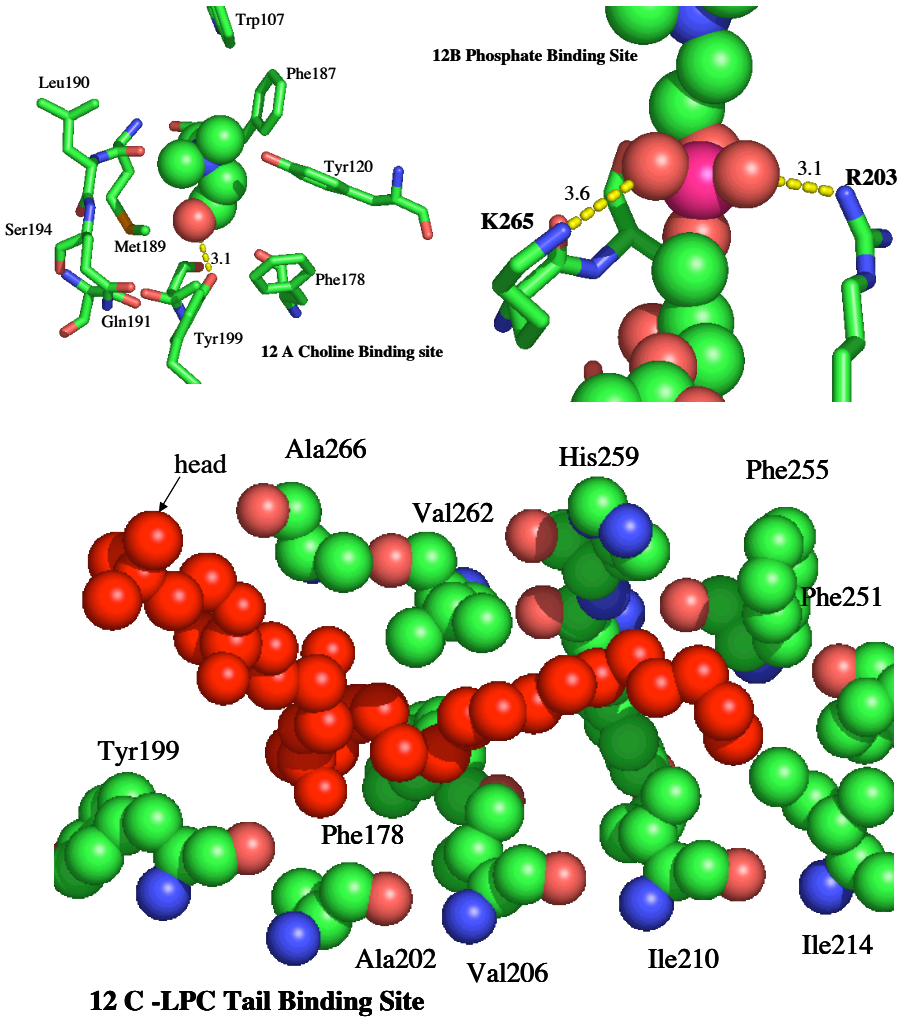

Fig. 12 A: Residues within 5A of the choline group of LPC in G2A receptor. **B:** residues within 5A of the phosphate group of LPC; **C:** Residues within 5A of the hydrophobic talk of LPC.

2.3 Predicted mutations that should decrease binding affinity of LPC to G2A receptor

Based on the predicted binding site of LPC in G2A we have identified the possible mutation candidates that would directly affect binding affinity of LPC to the mutant G2A. These mutants would be tested out in Witte's laboratory. Goddard group **thus predicted that the mutations, Arg203→Ala203 and Lys265→ Ala will reduce the binding to LPC significantly**. We first performed these mutations computationally on the receptor structure using the side chain replacement program, SCREAM. The calculated decrease in the binding affinity of LPC to the R203A mutant is by 7.65kcal/mol. The K265A mutant has a reduced affinity for LPC affinity by 4.35 kcal/mol. We also find that there are no hydrogen bonds made with the hydroxy group in LPC except for one with Lys265 (TM6).

2.6 Mutations that should increase LPC binding to G2A

The Goddard group also proposed that mutation Ala202→Ser and Ala266→ Ser will improve hydrogen bonds with LPC and thus enhance binding of LPC to G2A.

2.7 All possible mutations

Using theory we predict that the residues in the binding site of LPC that would affect binding are:

- Key Polar Residues: His174, Gln191, Tyr199, Arg203, Lys265.
- Key Non-Polar Residues: Phe123, Val206, Ile210, Pro211, Ile214, Phe255, Tyr258, His259, Val262.

3. Functional analysis of the R201A mutation predicted to decrease LPC binding to G2A

3.1. Generation of macrophage cell lines expressing wild-type and mutated G2A

For comparison purposes, the R201A mutation predicted to decrease binding of G2A to LPC was tested in parallel with several other mutants: N11Q should disrupt a putative N-Glycosylation site of unknown function at the N-terminus of G2A, the DRY motif mutant should impair coupling to G proteins and the L200S mutation should not decrease but enhance LPC binding. All the mutants have been generated using the QuickChange kit from Stratagene, sequenced to exclude the presence of secondary mutations and transferred into retroviral expression vectors (Clontech) using standard molecular cloning techniques.

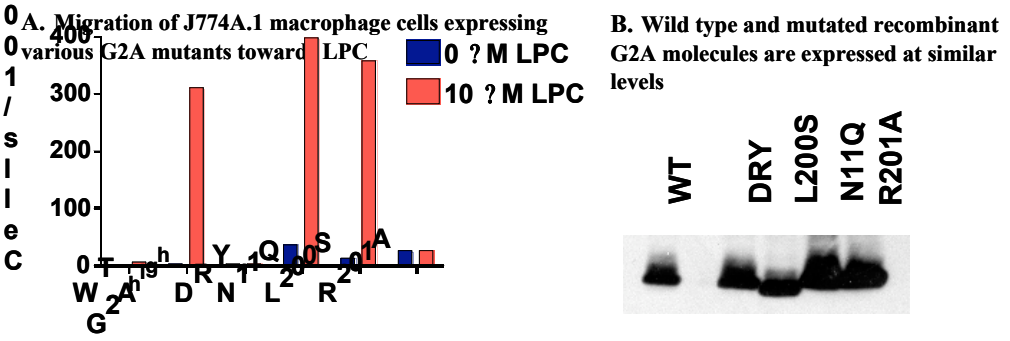

Fig. : (A) LPC -induced migration of J774 macrophages infected with G2A mutants. J774A.1 cells were transduced with retroviruses encoding wild type or mutated G2A and functio nal consequences of the mutations were assessed by examining cell mi gration towards LPC. The DRY motif and R201 were found to be critical for LPC -induced J774A.1 migration. WT: wildtype G2A; DRY: DRY motif mutation; N11Q: glycosylation site mutation; L200S and R201A: mutations of predicted LPC binding sites. **(B) Western blot using the rabbit polyclonal serum against G2A** .

3.2 Chemotaxis to LPC of macrophage cell lines engineered to express WT and mutated G2A

We have recently established a rapid functional assay for G2A based on the chemotactic responses of lymphoid and myeloid cell lines expressing this receptor to LPC [C.G. Radu, L.Yi, O.N.Witte, unpublished]. Briefly, J774A.1 cells [Ralph 1976] transduced with retroviruses encoding wild type G2A or the various mutants are added to the upper well of a Transwell cluster plate (Costar). A polycarbonate membrane with 5 mm pores separates the cells from the lower chamber to which LPC is added at a concentration (10 mM) previously shown to result in optimal chemotactic responses after a 2 hr incubation at 37°C in a 8% CO_2 incubator. At the end of the incubation period, cells that have transmigrated in response to LPC to the lower side of the membrane are fixed, stained

and counted under the microscope. Representative results from several chemotaxis experiments are shown in **Figure 6-9**.

In this assay, the DRY motif and the arginine residue at position 201 were found to be essential for G2A mediated chemotaxis to LPC. The N11Q and L200S mutants do not result in significant differences in migration compared to the wild type receptor. Quantitation of protein production by Western Blot (**Figure 6-9B**) shows equal amounts of R201A and wild type recombinant G2A and therefore excludes the possibility that this mutation actually affects protein folding. While the chemotaxis data supports the theoretical prediction, more experiments are required to conclude that R201 is actually involved in LPC binding.

Conclusions:

The joint theory/experimental approach to the study of the G2A receptor has resulted in an in-depth molecular understanding of the critical contacts for lipid binding by the orphan receptors. Further experiments are underway to test the remainder of the predicted mutations. Our studies have indicated essential contact points that stabilize the choline head, the phosphate portion, and the non-polar tail of the LPC lipid. Our suggested mutations and predicted active sites are likely to aid in the identification of the endogenous lipid for this system; modifications of the choline head residues are also likely to convert this LPC binding receptor to an LPA binding protein.

References:

- Berger EA, Murphy PM, Farber JM. (**1999**), Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol*; 17:657-700.

- Brenner S.E., Chothia C., Hubbard T.J.P., (**1998**), Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships, *Proc. Natl. Acad. Sci. USA,* 95, 6073-6078.

- Brinkmann V, Lynch K.R., (**2002**), FTY720: targeting G-protein-coupled receptors for sphingosine 1-phosphate in transplantation and autoimmunity. *Curr Opin Immunol*. Oct; 14(5):569-75.

- Chisolm, G., III, and Penn, M., (**1996**), Oxidized Lipoproteins and Atherosclerosis, Lippincott-Raven Publishers, Philadelphia, PA.

- Choi JW, Lee SY, Choi Y. (**1996**), Identification of a putative G protein-coupled receptor induced during activation-induced apoptosis of T cells. *Cell Immunol*. 168, 78-84.

- Chung S.Y., Subbiah S., (**1996**), A structural explanation for the twilight zone of protein sequence homology, *Structure* 4, 1123-1127.

- Connolly M.L., (**1983**) Analytical Molecular-Surface Calculation, *J. of Appl. Crystall.* 16, 548-558.

- Datta, D., Vaidehi, N., Floriano, W.B., Kim, K.S., Prasadarao, N.V. and Goddard III, W.A. (**2003**), Interaction of E-coli Outer membrane protein A with sugars on the receptors of the Brain Microvascular Endothelial Cells. *Proteins: Structure, Function and Genetics* 50, 213-221.

- Datta, D., Vaidehi, N., Xu, X., and W. A. Goddard III, (**2002**), Mechanism for antibody catalysis of the oxidation of water by singlet dioxygen, *Proc. Natl. Acad. Sci. USA*, 99, 2636.

- Datta, D., Vaidehi, N., Zhang, D., and Goddard III, W.A. **(2003 - submitted)**, Selectivity and Specificity of Substrate Binding in Methionyl-tRNA Synthetase, *Protein Science* , communicated.

- Ding, H.Q., Karasawa N., and Goddard III, W.A., Atomic Level Simulations on a Million Particles: The Cell Multipole Method for Coulomb and London Nonbond Interactions. *J. Chem. Phys*. **(1992)** 97:4309-4315.

- Eisenberg D., Weiss R.M., Terwilliger, T.C., The Hydrophobic Moment Detects Periodicity In Protein Hydrophobicity, (**1984**), *Proc. Natl. Acad. Sci. USA,* 8, 140-144.

- Ewing, T.A., and Kuntz, I.D. **(1997)**, Critical evaluation of search algorithms for automated molecular docking and database screening. *J Comput. Chem.* 18, 1175-1189.

- Floriano, W.B., Vaidehi, N., Zamanakos, G., and Goddard III, W.A., (**2003**), Virtual Ligand Screening of Large Molecule Databases using a hierarchical docking protocol (HierVLS), *J. Med. Chem*. accepted.

- Floriano, W.B., Vaidehi, N., And Goddard III, W.A. (**2003 - submitted**), Making Sense of Olfaction Through Molecular Structure And Function Prediction of Olfactory Receptors, *Chem. Senses.*

- Floriano, W.B., Vaidehi, N., Singer, M., Shepherd, G. and Goddard III, W.A., (**2000**), "Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor" *Proc. Natl Acad. Sci U.S.A.* , 97, 10712-10716.

- Freddolino, P.L., Kalani, M.Y., Vaidehi, N. Floriano, W.B., Trabanino, R.J., Freddolino, P.L., Kam, V. and Goddard III, W.A. (**2004**), Structure and function prediction for human 2- adrenergic receptor, PNAS.

- Fukushima N, Ishii I, Contos JJ, Weiner JA, Chun J., (**2001)** Lysophospholipid receptors. *Annu Rev Pharmacol Toxicol*; 41:507-34.

- Ghosh, A., Rapp, C.S., and Friesner, R.A., (**1998**), Generalized Born model based on a surface integral formulation. *J.Phys. Chem B* 102,10983-10990.

- Hall, S.E., Floriano, W.B., Vaidehi, N., Goddard III, W.A., **(2003-submitted)**, 3D Structures for mouse I7 and rat I7 olfactory receptors from theory and odor recognition profiles from theory and experiment, *Chem. Senses*, communicated.

- Hecht J.H, Weiner J.A, Post S.R, Chun J., (**1996**), Ventricular zone gene-1 (vzg-1) encodes a lysophosphatidic acid receptor expressed in neurogenic regions of the developing cerebral cortex, *J Cell Biol.*, 135(4): 1071-83.

- Huang MC, Graeler M, Shankar G, Spencer J, Goetzl EJ. (**2002**), Lysophospholipid mediators of immunity and neoplasia. *Biochim Biophys Acta.* May 23; 1582(1-3): 161-7.

- Im D.S., Heise C.E, Nguyen T, O'Dowd BF, Lynch K.R. (**2001**), Identification of a molecular target of psychosine and its role in globoid cell formation. *J Cell Biol.* 153, 429-34.

- Jain, A., Vaidehi, N., and Rodriguez, G. **(1993)**, " A Fast Recursive Algorithm for Molecular Dynamics Simulation", *J. Comp. Phys.* 106, pp 258-268.

- Jayasinghe S, Hristova K, White S.H., (**2001**), Energetics, stability, and prediction of transmembrane helices, *J. Mol. Biol*. 312 (5): 927-934.

- John B., and Sali A., **(2003)**, Comparative protein structure modeling by iterative alignment, model building and model assessment, *Nucleic Acids Research*, 31, 3982-

3992.

- Kabarowski J.H., Feramisco J.D., Le L.Q., Gu J.L, Luoh S.W., Simon M.I, Witte O.N. et al, (**2000**), Direct genetic demonstration of G alpha 13 coupling to the orphan G protein-coupled receptor G2A leading to RhoA-dependent actin rearrangement. *Proc Natl Acad Sci U S A*. Oct 24; 97(22): 12109-14.

- Kabarowski J.H., Xu Y., Witte O.N. (**2002**), Lysophosphatidylcholine as a ligand for immunoregulation. *Biochem Pharmacol*. 64(2): 161-7.

- Kabarowski JH, Zhu K, Le LQ, Witte ON, Xu Y. Lysophosphatidylcholine as a ligand for the immunoregulatory receptor G2A. (**2001**), *Science*. Jul 27; 293(5530):702-5.

- Kalani, Y., Vaidehi, N., Hall, S.E., Floriano, W.B., Trabanino, R.J., Freddolino, P.L., Kam, V., and Goddard III, W.A., (**2004**), Three-dimensional structure of the human D_2 dopamine receptor and the binding site and binding affinities for agonists and antagonists, *PNAS*. (Communicated).

- Kekenes-Huskey, P.M., Vaidehi, N., FlorianoW.B, and Goddard III, W.A., **(2003),** " Fidelity of phenyl alanyl tRNA synthetase"**,** *J. Phys. Chem* in print.

- Kim S.J., Gershov D., Ma X., Brot N., Elkon K.B., (**2002**), I-PLA(2) activation during apoptosis promotes the exposure of membrane lysophosphatidylcholine leading to binding by natural immunoglobulin M antibodies and complement activation. *J Exp Med*. 196(5): 655-665.

- Koh JS, Wang Z, Levine JS. (**2000**), Cytokine dysregulation induced by apoptotic cells is a shared characteristic of murine lupus. *J Immunol*. Oct 15; 165(8):4190-201.

- Lauber K, Bohn E, Krober S.M, Xiao YJ, Blumenthal S.G, Lindemann R.K, Marini P, Wiedig C., Zobywalski A, Baksh S, Xu Y, Autenrieth I.B, Schulze-Osthoff K, Belka C, Stuhler G, Wesselborg S. (**2003**), Apoptotic cells induce migration of phagocytes via caspase-3-mediated release of a lipid attraction signal. *Cell*. 113, 717-30.

- Le L.Q, Kabarowski J.H., Wong S., Nguyen K., Gambhir S.S., Witte O.N. (**2002**) Positron emission tomography imaging analysis of G2A as a negative modifier of lymphoid leukemogenesis initiated by the BCR-ABL oncogene. *Cancer Cell*. 1(4):381-91.

- Le L.Q., Kabarowski J.H., Weng Z., Satterthwaite A.B., Harvill E.T., Jensen E.R., Miller J.F., Witte O.N. (**2001**), Mice lacking the orphan G protein-coupled receptor G2A develop a late-onset autoimmune syndrome. *Immunity*. 14, 561-71.

- Liapakis G, Ballesteros J.A., Papachristou S, Chan W.C., Chen X, Javitch J.A., (**2000**), The forgotten serine - A critical role for Ser-203(5.42) in ligand binding to and activation of the beta(2)-adrenergic receptor, *J. Biol. Chem*. 275, 37779-37788.

- Lim, K-T, Brunett, S., Iotov, M., McClurg, R.B., Vaidehi, N., Dasgupta, S., Taylor, S. & Goddard III, W.A. (**1997**) Molecular Dynamics for Very Large Systems on Massively Parallel Computers: The MPSim Program. *J Comput. Chem*. 18, 501-521.

- Lin P, Ye R.D. (**2003**), The lysophospholipid receptor G2A activates a specific combination of G proteins and promotes apoptosis. *J Biol Chem*. 278, 14379-86.

- Lomize A.L., Pogozheva I.D., Mosberg H.I., (**1999**), Structural organization of G-protein-coupled receptors, *J. Computer-Aided Mol. Design* 13 (4): 325-353.

- Lusis AJ, (**2000**), Atherosclerosis. *Nature*. Sep 14; 407(6801): 233-41.

- MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck J.D., Field M.J., Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau F.T.K., Mattos C, Michnick S, Ngo T., Nguyen D.T., Prodhom B, Reiher W.E., Roux B., Schlenkrich M, Smith J.C, Stote R, Straub J, Watanabe, M., Wiorkiewicz-Kuczera J, Yin D, Karplus M, (**1998**), "All-atom empirical potential for molecular modeling and dynamics studies of proteins", *J. Phys. Chem*.B , 102, 3586-3616.

- Mayo, S. L., Olafson, B.D. & Goddard III, W.A. (**1990**), DREIDING - a generic force field for molecular simulations. *J. Phys. Chem*. 94, 8897-8909.

- McMurray HF, Parthasarathy S, Steinberg D. (**1993**), Oxidatively modified low density lipoprotein is a chemoattractant for human T lymphocytes. *J Clin Invest.* Aug; 92(2): 1004-8.

- Mitchison TJ. (**2001**), Psychosine, cytokinesis, and orphan receptors. Unexpected connections. *J. Cell Biol*. Apr 16; 153(2): F1-3.

- Nishimura H, Nose M, Hiai H, Minato N, Honjo T. (**1999**), Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity*. Aug; 11(2): 141-51.

- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller D.C., Okada T, Stenkamp R.E., Yamamoto M., Miyano M., **(2000)**, "Crystal structure of rhodopsin: A G protein-coupled receptor", *Science, 289*, 739-745.

- Parrill, A. L., Baker, D. L., Wang, D., Fischer, D. J., Bautista, D. L., van Brocklyn, J., Spiegel, S., and Tigyi, G. (**2000**) in *Lysophospholipids and Eicosanoids in Biology and Pathophysiology* (Goetzl, E. J., and Lynch, K. R., eds) Vol. 905, pp. 330–339, New York Academy of Sciences, New York.

- Quinn M.T., Parthasarathy S, Steinberg D. (**1988**), Lysophosphatidylcholine: a chemotactic factor for human monocytes and its potential role in atherogenesis. *Proc Natl Acad Sci U S A*. Apr; 85(8): 2805-9.

- Ralph P, et al. (**1976**), Lysozyme synthesis by established human and murine histiocytic lymphoma cell lines. *J. Exp. Med*. 143: 1528-1533.

- Rappé, A.K. & Goddard III, W.A. (**1991**), Charge Equilibration for Molecular Dynamics

Simulations. *J. Phys. Chem.* 95, 3358-3363.

- Rost B., (**1999**), Twilight zone of protein sequence alignments *Protein Engineering* 12, 85-94.

- Sali A, Potterton L, Yuan F, Vanvlijmen H, Karplus M., (**1995**), Evaluation of Comparative Protein Modeling by MODELER, *Proteins-Structure Function and Genetics* 23, 318-326.

- Schertler, G.F.X., Villa, C., and Henderson, R., (**1998**), Structure of rhodopsin. *Eye* 12:504-510.

- Shacham S, Topf M, Avisar N, Glaser F, Marantz Y, Bar-Haim S, Noiman S, Naor Z, and Becker O.M., (**2001**), Modeling the 3D structure of GPCRs from sequence, *Medicinal Research Reviews* 21 (5): 472-483.

- Shi L., Javitch J.A. (**2002**), The binding site of aminergic G protein-coupled receptors: The transmembrane segments and second extracellular loop *Ann. Rev. Pharm. Toxicology*, 42: 437-467.

- Strader, C.D., Fong, T. M., Tota, M.R., Underwood, D. and Dixon, R.A.F., (**1994**), Structure and Function of G-Protein Coupled Receptors, A*nnu. Rev. Biochem.* 63, 101-132.

- Sundstrom C, Nilsson K. (**1976**) Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int. J. Cancer* 17: 565-577.

- Tannor, D. J., Marten, B., Murphy, R., Friesner, R.A., Sitkoff, D., Nicholls, A., Ringnalda, M., Goddard III, W.A., and Honig, B., (**1994**)**,** Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from *ab initio* Quantum Mechanics and Continuum Dielectric Theory, *J. Am. Chem. Soc.* 116:11875-11882.

- Teller, D.C., Okada, T., Behnke, C.A., Palczewski, K., and Stenkamp, R.E., (**2001**), "Advances in determination of a high resolution three dimensional structure of rhodopsin, a model of G-Protein Coupled Receptor, *Biochemistry*, 40, 7761-7772.

- Trabanino, R.J., Hall, S.E., Vaidehi, N., Floriano, W.B., Kam, V., and Goddard III, W.A. (2004), First Principles Predictions of the Structure and Function of G-Protein Coupled Receptors: Validation for Bovine Rhodopsin, *BioPhys. J.,* (communicated).

- Vaidehi, N., Jain, A., and Goddard III. (**1996**) Constant Temperature Constrained Molecular Dynamics: The Newton-Euler Inverse Mass Operator Method. *J. Phys. Chem.* 100:10508.

- Vaidehi, N., Floriano W.B., Trabanino, R., Hall S.E., Freddolino, P. Choi, E.J. and Goddard III, W.A. (**2002**), "Structure and Function of GPCRs", *Proc. Natl. Acad. Sci., USA*, 99, 12622-12627.

- Vriend, G., (**1990**), "A molecular modeling and drug design program", *J. Mol. Graph.* 8, 52-56.

- Wang, P., Vaidehi, N., Tirrell, D.A. and Goddard, W.A. (**2002**), "Virtual ligand screening for phe analogs in phenylalanyl t-RNA synthetase", *J. Am. Chem. Soc*. 124, 14442-14449.

- Weng Z, Fluckiger A.C, Nisitani S, Wahl M.I, Le L.Q, Hunter C.A, Fernal A.A, Le Beau M.M, Witte O.N., (**1998**), A DNA damage and stress inducible G protein-coupled receptor blocks cells in G2/M. *Proc Natl Acad Sci U S A*. 95(21): 12334-9.

- White J, Haskins K.M., Marrack P., Kappler J. (**1983**), Use of I region-restricted, antigen-specific T cell hybridomas to produce idiotypically specific anti-receptor antibodies. *J Immunol*. 130, 1033-7.

- Xu Y, Zhu K, Hong G, Wu W, Baudhuin LM, Xiao Y, Damron D.S. (**2000**), Sphingosylphosphorylcholine is a ligand for ovarian cancer G-protein-coupled receptor 1, Nat Cell Biol. 2, 261-7.

- Zamanakos, G., Fast and Accurate Method for the Computation of Solvent effects in Molecular Simulations" Ph.D thesis, (**2001)**, Caltech, Pasadena.

- Zhang, D., N. Vaidehi, N. Goddard III, W.A. Danzer, J.F., and Debe, D. **(2002),** Structure-based design of mutant *Methanococcus jannaschii* tyrosyl-tRNA synthetase for incorporation of O-methyl-L-tyrosine, *Proc. Natl. Acad. Sci. USA*, Vol. 99, 6579-6584.

- Zohn IE, Klinger M, Karp X, Kirk H, Symons M, Chrzanowska-Wodnicka M, Der C.J, Kay R.J. (**2000**), G2A is an oncogenic G protein-coupled receptor. *Oncogene*, 19, 3866-77.