6-30-2006

Evolutionarily Conserved Allosteric Network In The Cys Loop Family Of Ligand-Gated Ion Channels Revealed By Statistical Covariance Analyses

Yonghui Chen

Kevin Reilly

Yongchang Chang
Barrow Neurological Institute, yongchang.chang@dignityhealth.org

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 26, pp. 18184–18192, June 30, 2006

Evolutionarily Conserved Allosteric Network in the Cys Loop Family of Ligand-gated Ion Channels Revealed by Statistical Covariance Analyses*

Received for publication, January 12, 2006, and in revised form, March 20, 2006 Published, JBC Papers in Press, April 4, 2006, DOI 10.1074/jbc.M600349200

Yonghui Chen[‡], **Kevin Reilly**[‡], **and Yongchang Chang**[§1]

From the [‡]*Department of Computer and Information Sciences, University of Alabama at Birmingham, Birmingham, Alabama 35294 and the* [§]*Division of Neurobiology, Barrow Neurological Institute, Phoenix, Arizona 85013*

The Cys loop family of ligand-gated ion channels mediate fast synaptic transmission for communication between neurons. They are allosteric proteins, in which binding of a neurotransmitter to its binding site in the extracellular amino-terminal domain triggers structural changes in distant transmembrane domains to open a channel for ion flow. Although the locations of binding site and channel gating machinery are well defined, the structural basis of the activation pathway coupling binding and channel opening remains to be determined. In this paper, by analyzing amino acid covariance in a multiple sequence alignment, we have identified an energetically interconnected network in the Cys loop family of ligand-gated ion channels. Statistical coupling and correlated mutational analyses along with clustering revealed a highly coupled cluster. Mapping the positions in the cluster onto a three-dimensional structural model demonstrated that these highly coupled positions form an interconnected network linking experimentally identified binding domains through the coupling region to the gating machinery. In addition, these highly coupled positions are also condensed in the transmembrane domains, which are a recent focus for the sites of action of many allosteric modulators. Thus, our results revealed a genetically interconnected network that potentially plays an important role in the allosteric activation and modulation of the Cys loop family of ligand-gated ion channels.

Ligand-gated ion channels (LGICs)[2] mediate fast synaptic transmission for communication between neurons. The Cys loop family of LGICs, with the signature cysteine loop in the amino-terminal domain, includes nicotinic acetylcholine receptors, serotonin receptor type 3, γ-aminobutyric acid receptor types A and C, glycine receptors, zinc activated cation channels, and invertebrate glutamate/serotonin-activated anionic channels or GABA-gated cation channels (1–4). Studies using site-directed mutagenesis, affinity labeling, cysteine accessibility test, and electron microscopy in the last two decades have demonstrated that all of the members of this receptor family have similar structural architecture (5). Each receptor is comprised of five subunits. Each subunit has a large amino-terminal extracellular domain that forms agonist-binding sites in subunit interfaces, four transmembrane domains (M1–M4) that form ion conduction pore, and a large intracellular loop that can interact with intracellular proteins for receptor targeting and regulation. The structure model of amino-terminal extracellular domain is further extended by the crystal structure of a homologous protein, acetylcholine-binding protein (6–9), and by the electron microscopic structure of the *Torpedo* nicotinic receptor (10). The structural model of nicotinic receptor transmembrane domains is also available via electron microscopy at 4 Å resolution (10, 11).

The Cys loop family of LGICs are allosteric proteins (3), in which binding of a neurotransmitter to its binding site in the extracellular amino-terminal domain controls distant gating machinery in the transmembrane domain to open the ion conduction pore. The kinetic mechanism of channel activation can be best described by an allosteric model in which agonist binding and channel gating are highly coupled (3, 12, 13). This long range coupling of the agonist-binding domain to the gating machinery requires an interconnected allosteric network, through which binding energy can be reliably transmitted, in the form of a "conformational wave" (14), from the agonist-binding site to the gating machinery to open the channel. Information about this interconnected allosteric network, however, is not readily available by directly examining the structural models. Although recent experimental studies have made significant contributions toward understanding the mechanism of ligand-gated ion channel activation (4, 15), exhaustive experimental studies are time consuming, and the activation pathway still needs to be defined. Thus, to facilitate future experimental studies, it is necessary to use computer-aided analysis to define the entire allosteric network for experimental validation.

Statistical coupling analysis (SCA) is a sequence-based statistical method designed to estimate the thermodynamic coupling of two residues in a protein. The basis of this method is that the coupling of two sites in a protein, either directly or allosterically, should cause these two positions to coevolve. Such coevolved residues can be identified by analyzing a large and diverse multiple sequence alignment (MSA) of a protein family for the distribution probability of 20 amino acid residues at each position (16). With this method, the degree of residue covariance at two sites, in a form of "coupling energy," can be determined by observing the effect of perturbation at one site (extracting a subset of sequence alignment containing a relatively conserved residue at the site) on the amino acid distribution of another site. Prediction of potential interacting residues could dramatically reduce the work of exhaustive mutagenesis scanning and facilitates identification of functionally important residues in the interconnected allosteric network of the protein for the mechanisms of binding-gating coupling of the entire family. This method has been successfully used to define interconnected allosteric networks of several protein families, such as PDZ domains (16), G-proteins (17), G-protein-coupled receptors, serine proteases, globins (18), and retinoid X receptors (19).

* This work was supported by funds from the Barrow Neurological Foundation and Women's Board (to Y. Chang). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed: Division of Neurobiology, Barrow Neurological Institute, 350 West Thomas Rd., Phoenix, AZ 85013. Tel.: 602-406-6192; Fax: 602-406-4172; E-mail: yongchang.chang@chw.edu.

[2] The abbreviations used are: LGIC, ligand-gated ion channel; GABA, γ-aminobutyric acid; GABAR, GABA receptor; MSA, multiple sequence alignment; SCA, statistical coupling analysis; McBASC, McLachlan-based substitution correlation; M1–M4, transmembrane domains 1–4.

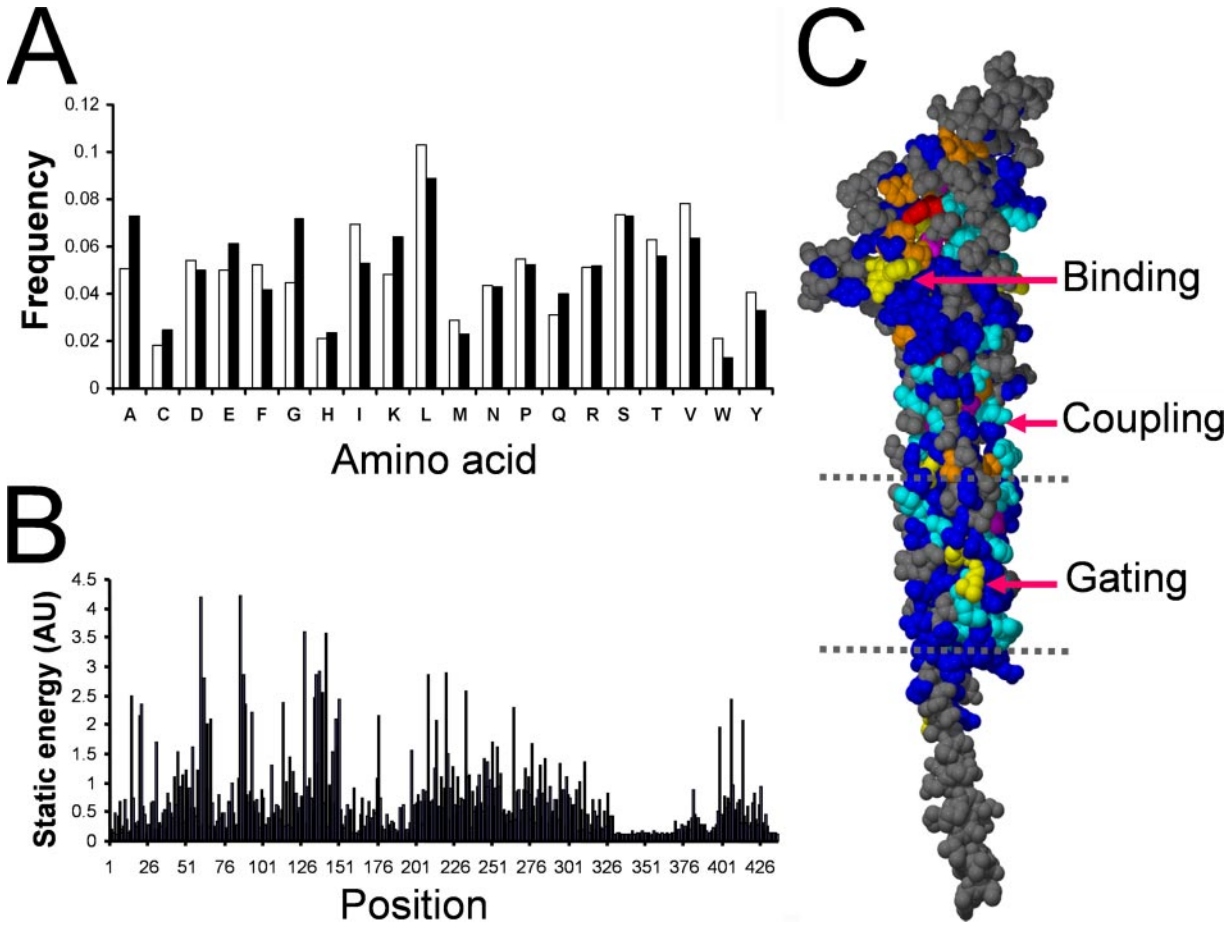

FIGURE 1. **Amino acid frequencies and static energy.** *A*, amino acid frequencies in all proteins (*filled bars*) or in the MSA (*open bars*). *B*, static energy (arbitrary unit) in all positions using the numbering of the *Torpedo* nicotinic receptor α subunit. *C*, mapping of static energy onto the structure of the *Torpedo* nicotinic receptor α subunit. *Gray*, <0.5; *blue*, 0.5 to <1; *cyan*, 1 to <1.5; *yellow*, 1.5 to <2; *orange*, 2 to <2.5; *purple*, 2.5 to <3.0; *red*, ≥3. Two *dotted lines* represent the two surfaces of the plasma membrane.

McLachlan-based substitution correlation (McBASC) is another approach to find covariant positions in a protein family (20), although it is more frequently used to find direct contacting residues (21). By comparing pairs of sequences in an MSA, this method assigns a score for each comparison at each position based on the change of amino acid residue properties using the McLachlan substitution matrix (22). Correlation analysis (correlation coefficient) of these mutational scores between two sites from the MSA of a protein family then can be used to identify coevolved sites.

In this paper, using these two approaches with different scoring systems, we have identified a cluster of genetically covariant sites in the Cys loop receptors. Mapping these positions onto the three-dimensional structural model of a nicotinic receptor subunit reveals that these positions are mainly clustered in functionally important domains, forming an interconnected allosteric network linking the agonist-binding pocket to the gating machinery via coupling domains. In addition, these highly coupled positions are also clustered in transmembrane domains, the recent focus for the sites of action of many allosteric modulators. Thus, our results revealed a genetically interconnected network that potentially serves as the activation pathway and plays an important role in allosteric modulation of the Cys loop family of LGICs.

EXPERIMENTAL PROCEDURES

Data Source and Multiple Sequence Alignment—The amino acid sequences of subunits in the Cys loop receptor family of ligand-gated ion channels were downloaded from the Ligand-Gated Ion Channel Data base in the European Bioinformatics Institute website (www.ebi.ac.uk/compneur-srv/LGICdb/LGIC.html), where a redundancy check has been performed. Based on the length distribution histogram of all of the sequences (data not shown), we excluded those sequences that clearly do not belong to the same population. Extra long sequences (>700 residues) could have different structure, and extra short sequences (<250 residues) are likely incomplete sequences that would introduce unnatural gaps and influence coupling analysis (see Fig. 2*B*). Thus, these extra long and short sequences were excluded for further analysis. The remaining 389 sequences were used for analysis. All of the sequences were aligned using the Clustalw1.83 package with default parameters: 10.00 gap opening penalty, 0.20 gap extension penalty, and Gonnet series of the protein weight matrix. Because the structural model of the *Torpedo* nicotinic receptor is the best model available, for all calculations, the numbering in the α subunit of *Torpedo california* nicotinic receptor was used, ignoring the signal peptide and gaps inserted into the subunit during the sequence alignment.

Statistical Coupling Analysis (SCA)—The static energy ($\Delta G_i^{\mathrm{stat}}$) for each site *i* and coupling energy ($\Delta\Delta G_{i,j}^{\mathrm{stat}}$) between two sites, *i* and *j*, were calculated following Lockless and Ranganathan (16) using the software written in JAVA adapted from the original software written in C (kindly provided by Dr. Rama Ranganathan at University of Texas Southwestern Medical Center at Dallas).

Correlated Mutational Analysis—Correlated mutational analysis was carried out using the McBASC (23). The program for this calculation, written in JAVA, was downloaded from Anthony A. Fodor's website

(www.afodor.net), modified for formatted output, and executed under a JAVA environment.

Clustering Analysis—To extract information from the large data sets for coupling or correlated mutation analysis, a clustering analysis was performed using Hierarchical Clustering Explorer 3.0 by Jinwook Seo at University of Maryland (www.cs.umd.edu/hcil/multi-cluster/hce3.html). The coupling energy/correlation coefficient data matrices without normalization were clustered with complete linkage.

Visual Presentation—For visual presentation of the highly coupled residues in the structural model of a subunit, the structure of the α subunit of the *Torpedo marmorata* nicotinic receptor was downloaded from the National Center for Biotechnology Information website with the Protein Data Bank identification code of 2BG9. The sequence of this subunit is incomplete and therefore not included in covariance calculation. However, it is highly homologous to the α subunit of *T. california* nicotinic receptor, and both subunits have identical numbering. Thus, the sites with high coupling energy or correlation coefficient were directly mapped onto the above structural model without further conversion, using the molecular graphics program DEEPVIEW/SwisspdbViewer v3.7 (us.expasy.org/spdbv/). The resulting image was saved as a POV-Ray 3.5 scene file. The final image of the model was rendered by POV-Ray 3.6 software (www.povray.org/download/). Visualization of the coupling energy result before clustering (Fig. 2*A*) was accomplished by using TreeView version 1.60 (rana.lbl.gov/EisenSoftware.htm).

RESULTS

Static Energy—To calculate the coupling, we started by counting occurrences of amino acids at each position in the MSA. Fig. 1*A* shows the relative frequencies of the amino acid residues in the Cys loop family of LGICs (*open bars*) and in all proteins from the Swiss-Prot data base (*filled bars*) used for calculation. Note that the frequencies of amino acid residues in the MSA slightly deviate from those in all proteins. Hydrophobic residues, such as Leu, Ile, Phe, Val, and Trp in the MSA, had slightly higher frequencies than average. Small and some hydrophilic residues such as Gly, Ala, and Lys had lower frequencies than average. This is expected for membrane proteins with multiple hydrophobic transmembrane stretches. The amino acid frequencies at each site were then determined and converted to probabilities for all 20 amino acids (16). The probabilities then were used to calculate the static energy. If the amino acid distribution at a site is similar to the distribution for all positions, then the site is not conserved, and the static energy approaches zero. In contrast, if a site is conserved, its amino acid distribution will deviate from the mean, and the static energy at that site will be higher. Thus, the magnitude of the static energy represents the extent of deviation of the amino acid distribution at each site from the mean in the MSA and therefore represents the extent of residue conservation at that site.

Fig. 1*B* shows the static energy for all 437 positions using the numbering of *T. california* nicotinic receptor α subunit. Note that a stretch of positions toward the carboxyl terminus had low static energy. This region corresponds to the large intracellular loop between M3 and M4, the most diversified region in this protein family. The static energy for all positions was then mapped onto the structural model of the *Torpedo* nicotinic receptor α subunit (Fig. 1*C*). The most conserved positions (in *red*) are mainly located in the protein core near the binding site, and intermediately conserved positions are clustered in functionally important domains for binding, coupling, and channel gating as indicated by *arrows*.

Coupling Energy—To calculate the statistical coupling energy, we performed perturbation analysis as described by Suel *et al.* (18). Briefly, sequences containing a conserved residue (>30%) at a particular site were taken out of the MSA to form a subset. There are 253 sites with at least one relatively conserved residue (>30%). Thus, 253 perturbations were performed (one perturbation at each site), and 253 subsets were generated. The extracted sequences in a subset containing only the conserved residue at the perturbation site resulted in amino acid redistribution at this and all the other sites. The amino acid probabilities at each site in a subset were then determined and used for coupling energy calculation. If the perturbation at one site significantly changes the amino acid distribution at another site, then these two sites have high coupling energy. Otherwise, they have low coupling energy.

The calculation resulted in a 437 × 253 matrix of the coupling energy (Fig. 2*A*). In some regions of the receptor, such as the large intracellular loop (sites 301–402) between the third and fourth transmembrane domains with the most diversified sequences and low static energy (Fig. 1*B*), the alignment generated large gaps at many positions. To determine whether gaps can influence the coupling result, we examined the relationship between the number of gaps at each position and the mean coupling energy of all positions in response to the same perturbation. Fig. 2*B* plots the number of gaps against the mean coupling energy for each perturbation. Note that all of the positions with more than 60 gaps had high mean coupling, suggesting the number of gaps does have some influence in the coupling energy calculation. To avoid this potential influence, we discarded positions with more than 60 gaps and most of the M3-M4 intracellular loop (sites 296–392) for further analysis in both rows (coupling) and columns (perturbation). This resulted in a 311 × 219 matrix.

To identify highly coupled sites from this large data set, we performed a clustering analysis. Fig. 3*A* shows the clustering result of coupling energy for this matrix with 219 rows (perturbation) and 311 columns. Note that the sites with high coupling energy are mainly clustered in the *bottom right* as indicated by the *three yellow boxes*. Fig. 3*B* is a closer view of these clusters. The positions of all of the columns in this highly coupled cluster showed a similar coupling pattern to many perturbations, suggesting that they are covariant in response to same set of perturbations and thus are mutually coupled. The detailed positions in Fig. 3*B* are listed in Table 1.

Correlated Mutation Analysis—With the same set of MSA, we performed correlated mutation analysis using the McBASC method. This resulted in a 437 × 437 matrix (data not shown). Similarly, to extract information from the large data set, we first removed positions with large gaps (>60) and intracellular loop (sites 296–392) to avoid potential influence of gaps and improper alignment. The remaining data were clustered using the Hierarchical Clustering Explorer 3.0 software. The results are shown in Fig. 4*A*. Note that there is a high correlation coefficient cluster (the *yellow box* in the *bottom right corner*) from the large background. The details of this cluster with high correlation coefficient are shown in Fig. 4*B*, and the positions in this cluster are listed in Table 1.

DISCUSSION

In search for the activation pathway, we used two statistical analyses along with clustering to systematically identify the genetically interconnected positions in the Cys loop family of ligand-gated ion channels. Highly coupled positions predicted by both methods overlapped by nearly 70% (see below). Mapping these positions onto the three-dimensional structural model demonstrated that these highly coupled positions were mainly clustered in important functional domains, linking

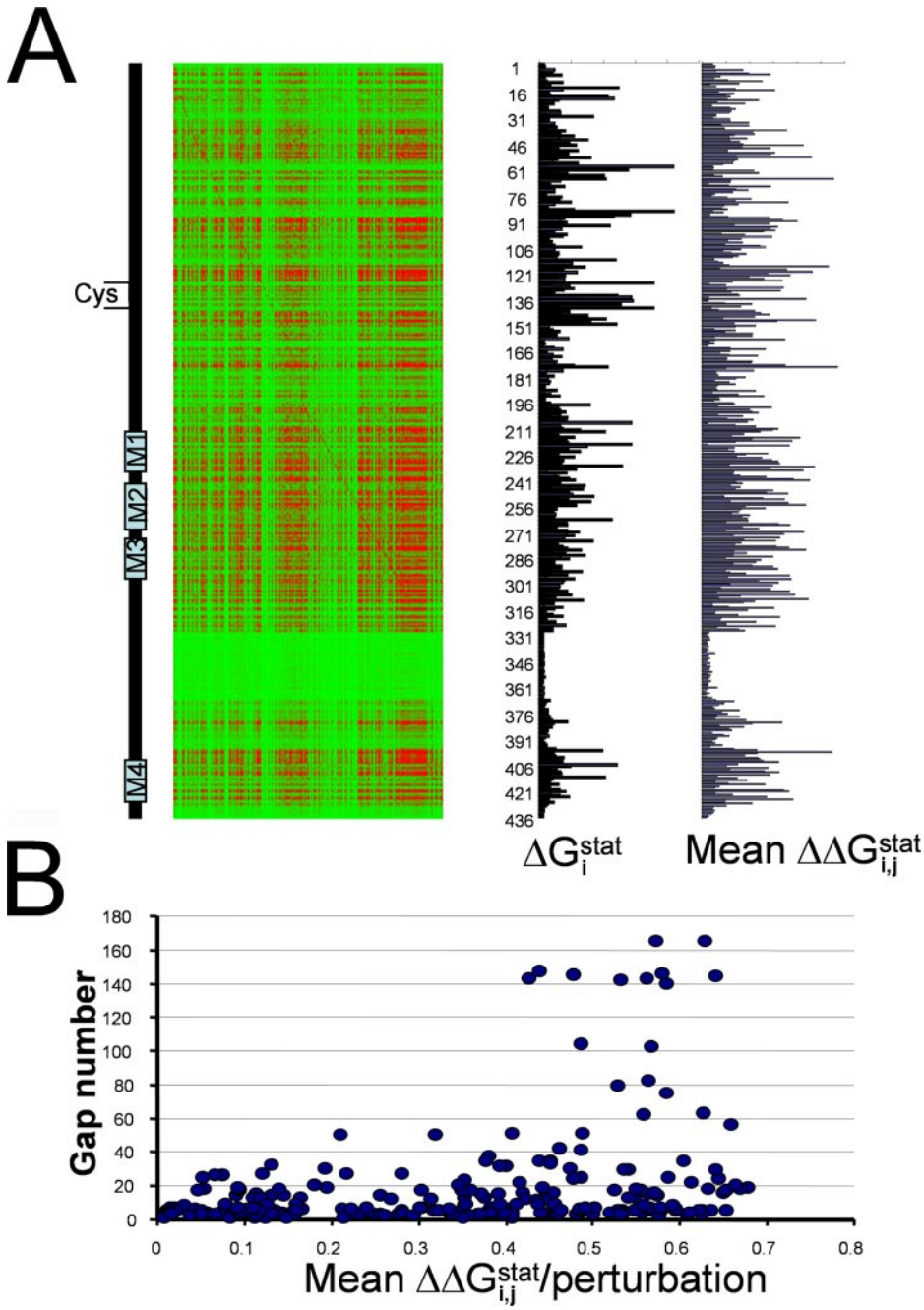

FIGURE 2. **Coupling energy and its relationship to the inserted gaps during alignment.** *A*, coupling energy in all positions (*rows*) with 253 different perturbations (*columns*). Cysteine-loop and four transmembrane domains are shown on the *left*. The static energy (ΔG_i^{stat}) and mean coupling energy (mean $\Delta\Delta G_{i,j}^{stat}$) for all perturbations are plotted in corresponding positions. *B*, relationship between mean coupling energy per perturbation and number of gaps at each site.

the binding pocket through coupling domains to the gating machinery. Thus, our results suggest an interconnected network that may serve as the allosteric activation pathway, coupling agonist binding to channel function. The finding can be used as a guide for experimental design and to facilitate elucidation of the activation mechanism for the Cys loop family of LGICs.

Comparison of Coupling and Correlation Results—To compare the identified covariant positions by the two methods, we list these positions in Table 1. Note that positions predicted by the two methods substantially overlap. In fact, the overlapping sites represent 62% of the total number of positions predicted by SCA and 65% of the total number of positions predicted by McBASC. If we take an additional stringent step by removing the sites with 20 or more gaps (sites 7, 11, 22, 81, 95, 166, 230, 240, 398, 399, 425, and 429), then the results are more consistent, and the overlapping sites represent 69 and 68% for the predictions by SCA and McBASC, respectively. The prediction differences could be due to different scoring methods: SCA uses amino acid probability and observes changes in the probability distribution in response to a perturbation at a site by extracting a fraction of total number of sequences containing a relatively conserved residue, whereas McBASC uses a score matrix with consideration of amino acid properties and compares all possible pairs. Thus, theoretically McBASC more effectively uses sequence data and therefore could be a better predictor for genetically covariant positions in a protein family. Nevertheless, the positions predicted by both methods are the most reliable ones with high coupling. Positions predicted by only one method still should be coupled but with slightly lower coupling strength.

To visualize this genetically interconnected network, we mapped the positions predicted by both methods (after removing the sites with 20 or more gaps) in the three-dimensional structure of the *Torpedo* nicotinic

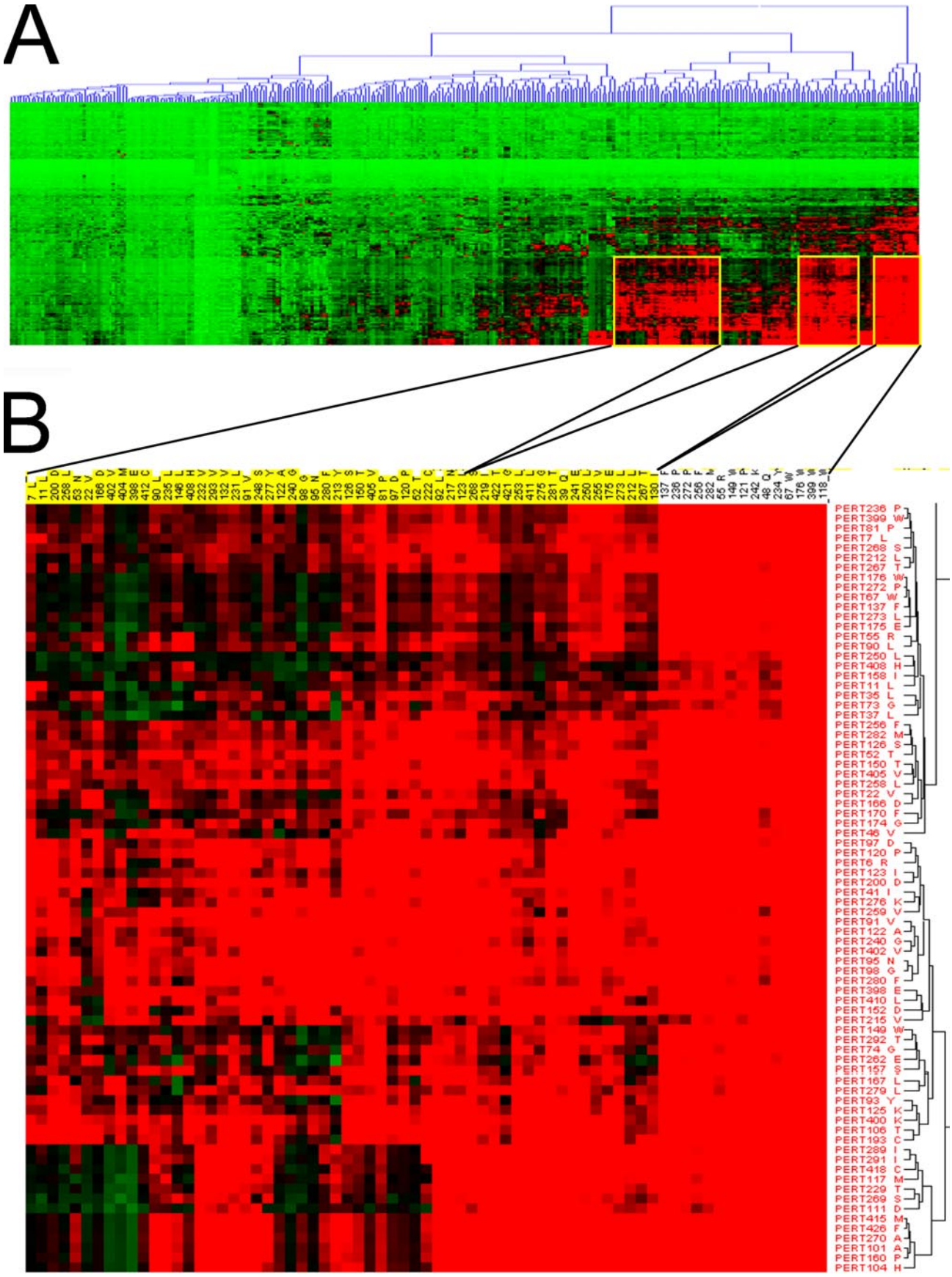

FIGURE 3. **Coupling energy for all pairs excluding sites with 60 gaps and most sites of the intracellular loop.** *A*, clustering results of the coupling energy. Highly coupled clusters are highlighted in *yellow boxes*. *B*, closer view of these highly coupled positions. The detailed sites in this cluster are listed in Table 1.

receptor α subunit (Fig. 5*A*). Each position in the high coupling cluster is shown with its side chain and color-coded as follows: *red* represents the positions predicted only by McBASC; *green* represents positions predicted only by SCA; and *yellow* represents positions predicted by both methods. Note that *yellow residues* form a sparse network with high density in binding domains, coupling region, and gating machinery. The *red* and *green residues* further fill gaps, forming an interconnected network.

Two salient features are apparent in this interconnected network. First, they are highly clustered in functionally important domains connecting the agonist-binding site through the coupling region to the gating machinery, forming a putative activation pathway. Second, many positions are concentrated in a region of the recent focus for the sites of action of many allosteric modulators (24): transmembrane domains. These two aspects are further discussed in detail below.

TABLE 1
Highly coupled sites identified by SCA or McBASC

SCA	McBASC	SCA	McBASC	SCA	McBASC	SCA	McBASC
7		121	121		226		276
11		122	122		230	277	
22		123	123	231		280	280
39			124	232	232	281	281
	41	126	126	234		282	
48	48	130		235			286
	49	132	132	236	236		291
	50	137	137	240	240		292
52	52	146		241	241	293	293
53		149		242	242	398	398
55		150	150	248		399	399
67	67	166			249	402	402
81	81		170	250	250	404	
90			173	253		405	405
91	91	175		255	255	408	
92	92	176	176	256	256	410	
95	95	200	200	258	258	411	411
	96		203		259	412	
97			211	267	267		415
98	98	212	212	268	268	421	421
	102	213			270	422	422
	112	217		272	272		425
118		219	219	273	273		429
120	120	222		275	275		

Allosteric Activation Pathway—The strategic location of the highly coupled residues strongly suggests their importance in channel function. The agonist-binding pocket of a receptor is located in the amino-terminal domain at a interface between two subunits, each contributing three binding loops (7). Fig. 5 (*B* and *C*) plots highly coupled residues in the context of other residues in two different views (principal face and complementary face). The residues in the high coupling cluster for all three colors in Fig. 5*A* are now in *yellow*. Important binding sites are highlighted with *red* in the principal face and *cyan* in the complementary face. For the convenience of numbering, both faces are shown in the same subunit. In reality, this is only true for homomeric channels in that the same subunit contributes both the principal face and the complementary face of the binding sites. In heteromeric channels, the principal and complementary faces of the binding pocket are in different subunits. The overlapping residues are in *orange* for the overlapping between *red* and *yellow* or in *green* for overlapping between *cyan* and *yellow*. The gate forming residue, the conserved M2 Leu, is highlighted in *purple*. Note that only two highly coupled positions (sites 55 and 149) overlap with binding site residues. This is because many functionally important residues are highly conserved and nonvariant and thus escape detection by covariant analysis. However, these binding site residues are flanked by highly coupled residues in both the principal and complementary faces (Fig. 5, *B* and *C*). With the exception that predicted high coupling positions flank binding residues in loop E from the top of the molecule (Fig. 5*C*, sites 67 and 112), all of the other positions form an interconnected network connecting the binding pocket through the coupling region (see below) to the gating machinery. Interestingly, in the amino-terminal domain, the highly coupled residues are distributed only in the inner sheet. This is consistent with current understanding of the activation mechanism as suggested by 4 Å electron microscopic study: activation involves a

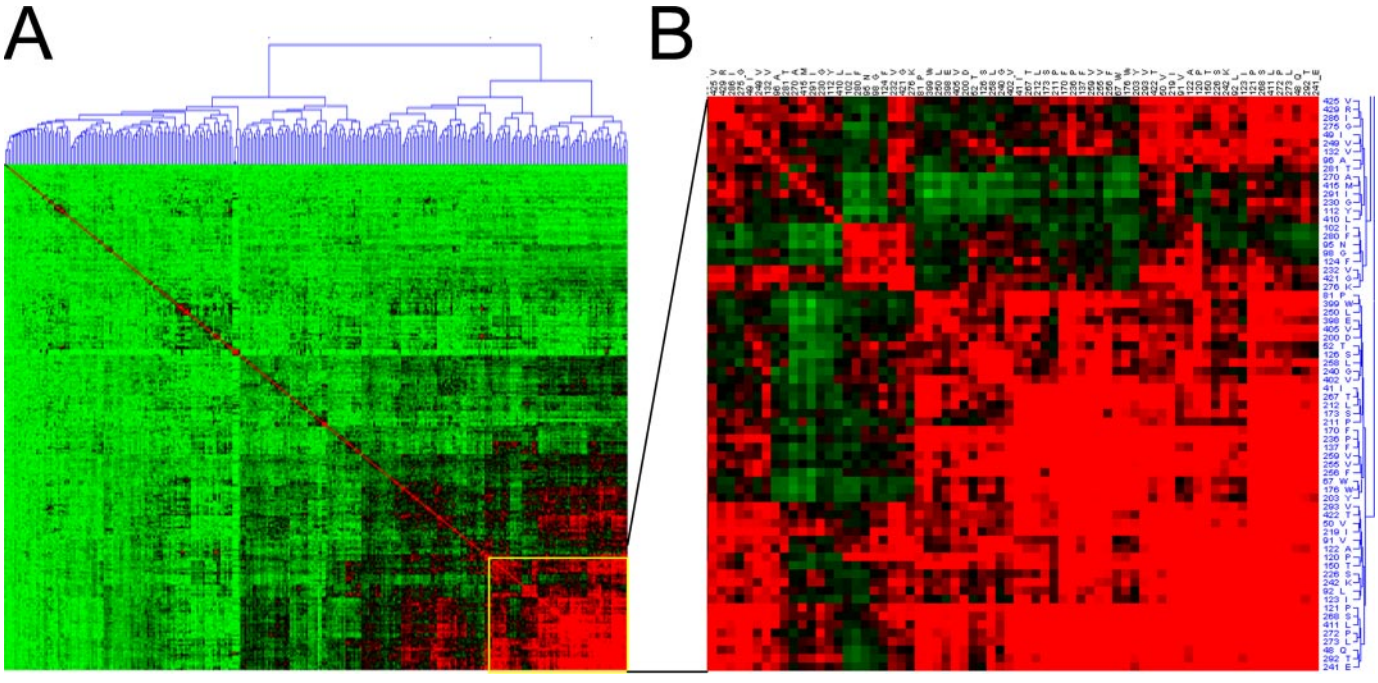

FIGURE 4. **Correlation coefficient for all pairs excluding sites with 60 gaps and most sites of the intracellular loop.** *A*, clustering results of the correlation coefficients. The cluster with high correlation coefficient is highlighted in a *yellow box*. *B*, close view of the highly correlated cluster. The detailed sites in this cluster are listed in Table 1.

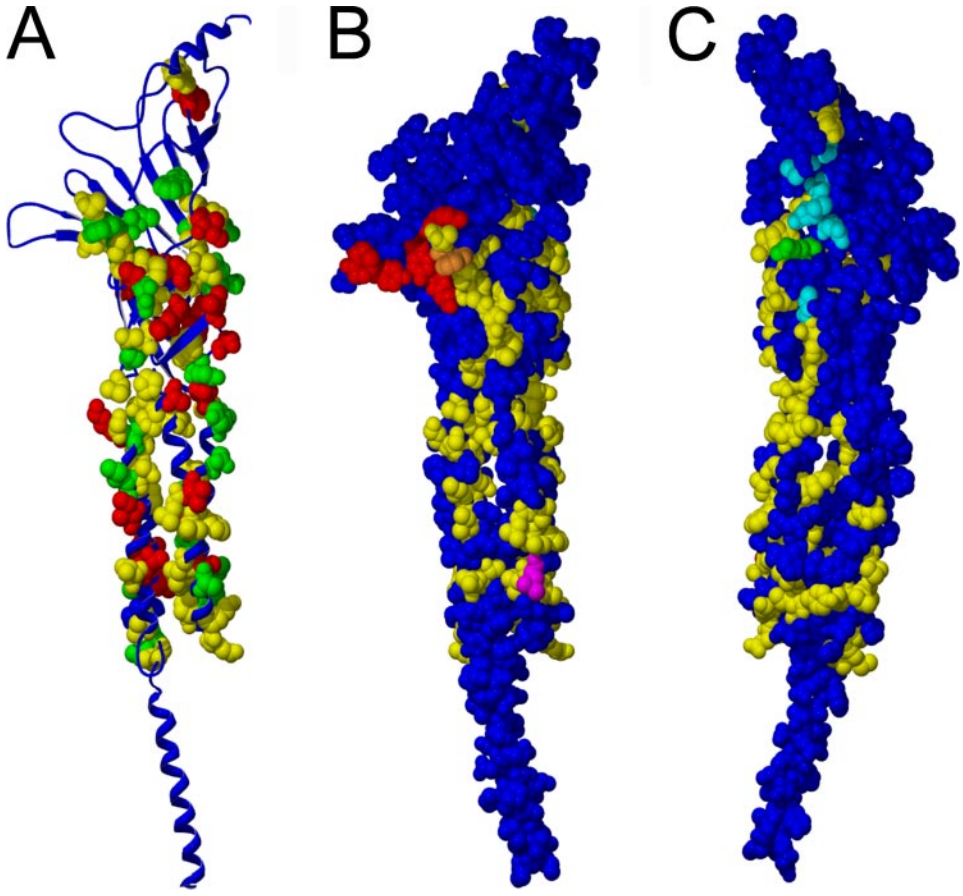

FIGURE 5. **Mapping highly coupled sites onto the three-dimensional structure of the *Torpedo* nicotinic receptor α subunit.** *A*, covariant sites predicted by SCA only (*green*), by McBASC only (*red*), or by both methods (*yellow*). *B*, principal face of the binding pocket (*red*) and other regions with all predicted coupling sites (*yellow*) and gate forming M2 leucine (*purple*). The *orange residue* represents overlap between *red* and *yellow*. *C*, complementary face of the binding pocket (*cyan*) and other regions with all predicted coupling sites (*yellow*). The *green residue* represents the overlap between *cyan* and *yellow*. All of the binding residues are from Fig. 1 in the paper by Brejc *et al.* (7).

clockwise rotation of the inner sheet of the amino-terminal domain around its own axis in each subunit (10, 11).

The highly coupled positions also clustered in the contact region between the amino-terminal domain and transmembrane domain. This is a region that is believed to be crucial in coupling binding to channel gating. In fact, the coupling between amino-terminal domain and channel domain has been postulated to be mediated by the M2-M3 linker (25–28). More recent studies with electron microscopy structure of nicotinic receptor (10, 11) or mutagenesis studies (29–32) further suggest that it is mediated by interactions between amino-terminal domain loop2/loop 7 and transmembrane domain linker M2-M3, although crucial residues involved in coupling vary with different receptors. The rate-equilibrium free energy relationship analysis suggests that both loop 2 and loop 7 (cysteine loop) are involved in channel activation (33). Loop 9 (Loop F) is also required for the function of a chimera channel (34). More recently, mutant cycle analysis in nicotinic acetylcholine receptor (35) or unnatural amino acid substitution in serotonin receptor type 3 (36) have identified a key residue in the M2-M3 linker for channel activation. This functionally important residue does overlap with the high coupling site 272. Moreover, the residues required for benzodiazepine allosteric coupling in M2 and M2-M3 linker (GABARγ2T281,I282,S291 (37) and GABARα1V279 (38)) overlap with the highly coupled positions. In addition, our results also provided a potential link between the binding pocket and the coupling region (loops 2 and 7) in this putative allosteric network (Fig. 5*A*), which may represent the physical basis for inner sheet movement as a "rigid body" during channel activation (11).

Finally, highly coupled residues are also clustered in the middle and intracellular end of the M2 domain, a region with the putative ion channel gate (Fig. 5*B*, the *purple residue* in the transmembrane domain) as suggested by many studies ((1, 5, 13, 39–41) and ultimately confirmed by electron microscopy studies of nicotinic receptor at 4Å resolution (10, 11). Again, the conserved M2 leucine is not predicted by covariant analysis but is surrounded by highly coupled residues. The highly coupled positions, however, do cover the region in the beginning of M2 in the intracellular end, the location of the selectivity filter, which differentiates cationic nicotinic and serotonin receptor channels from anionic GABA- and glycine receptor channels (42–46). In addition, other residues in the M2 domain (47, 48) and other transmembrane domains such as pre-M1 (49), M1 (50, 51), M3 (52), and M4 (53, 54) are also important in channel gating, and the M1-M2 and M2-M3 linkers have been suggested to act as hinges governing allosteric control of the M2 domain (11, 27). Given the significance of all four transmembrane domains in channel gating, it is understandable that the highly coupled cluster covers these transmembrane domains.

In summary, we have identified an interconnected network that physically links agonist-binding domains to channel gating domain. This would represent the entire allosteric network, through which binding signals in the amino-terminal domain can be transduced to gating function in the distant location.

Sites of Action for Allosteric Modulators—In addition to the gate-containing M2, our results showed that the highly covariant cluster also includes positions in the M1, M3, and M4 domains. All four transmembrane domains, especially the extracellular half of M2 and M3, are recently recognized as important sites of action for many allosteric modulators such as alcohol, general anesthetics, neurosteroids, and barbiturates (24, 55). Allosteric modulators for ligand-gated ion channels are compounds binding to a site distinct from the agonist-binding site. With the exception of benzodiazepines, which, like agonists, bind to the amino-terminal domain but in a different subunit interface, most allosteric modulators exert their

action by binding to the transmembrane domains of the receptor. In fact, the sites of action for many allosteric modulators in all four transmembrane domains overlap or flank the highly coupled positions. These include a site for barbiturate/neurosteroid/etomidate/propofol modulation (GABARβ2G219) (56) in M1; sites for alcohol (GlyRα1S267) (57), volatile anesthetics (GABARα1S270) (58) or neurosteroid (GABARρ1I307) (59) in M2; sites for alcohol/anesthetics (GlyRα1A288) (57), barbiturates (GABARρ1W328) (60), and redox (GABARβ3C313) (61) in M3; and sites for alcohol binding (nAChRαH408, C412) (62) in M4. By binding to the transmembrane domains, these allosteric modulators can alter the energy landscape to favor channel opening and potentiate neurotransmitter action. Many of them, such as barbiturates, neurosteroids, and general anesthetics, can even selectively open channels directly. Again, some functionally important residues in channel gating and modulation in this region can be highly conserved and escape detection by covariant analysis. Nevertheless, given that many sites are overlapping with or flanking the experimentally identified sites for the action of many allosteric modulators, we have enough reason to believe that the predicted interconnected allosteric network should also serve as the framework to mediate allosteric modulation of this receptor family.

Concluding Remarks—Our finding that the highly coupled cluster spans the regions from the binding pocket to the gating machinery re-emphasizes an important concept: binding and channel function are mutually coupled (12, 63). This long range coupling requires an interconnected allosteric network. Perturbation of this allosteric network, either by agonist binding or mutations in binding domains (*e.g.* loops A (64, 65), B (66), D (67), or E (64)) or gating machinery (13, 40, 41, 47, 48, 68) can alter channel gating behavior and even make a channel open spontaneously in the absence of agonist. Thus, it is the fine balance of all residues in this allosteric network that determine the function of the channel, from agonist binding to channel gating. Fine tuning of this allosteric network with coordinated changes of the side chains of amino acid residues during long evolution preserves channel function and generates functional diversity of the channels in this family to meet the growing need of ever evolving brain function.

Although our results can provide a useful general reference for structural dynamics studies of ligand-gated ion channels, caution should be exercised when applying the results to a particular member of the Cys loop family at precise positions. First, because of the nature of the interconnection with coordinated mutations, the effect of single point mutation at a particular site on channel function may vary with different receptors. Second, our analysis could be limited to the coupling between residues within one subunit. It may not account for the interaction between subunits. Because receptors in the Cys loop receptors have a pentameric structure with five subunits in a receptor, interactions between subunits are also important for receptor structure and function. Although detailed interaction between subunits can be determined by another type of analysis such as the subtractive correlated mutation method analyzing linked subunits (69), positions for this intersubunit interaction may be already embedded in the covariant sites of our results, because our analysis includes all subunits. Furthermore, the conformational change in the amino-terminal domain is proposed to be coupled to the channel gating machinery within each subunit (11). Thus, our results can still provide valuable information for the mechanisms of activation and modulation for the Cys loop receptors.

Acknowledgments—We thank Dr. Rama Ranganathan (University of Texas Southwestern Medical Center at Dallas) for providing the method for statistical coupling analysis and Dr. Anthony A. Fodor (Stanford University) for providing the free download program for McBASC analysis.

REFERENCES

1. Unwin, N. (1993) *J. Mol. Biol.* **229,** 1101–1124
2. Ortells, M., and Lunt, G. (1995) *Trends Neurosci.* **18,** 121–127
3. Changeux, J. P., and Edelstein, S. (1998) *Neuron* **21,** 959–980
4. Lester, H., Dibas, M., Dahan, D., Leite, J., and Dougherty, D. (2004) *Trends Neurosci.* **27,** 329–336
5. Karlin, A., and Akabas, M. (1995) *Neuron* **15,** 1231–1244
6. Celie, P., Klaassen, R., van Rossum-Fikkert, S., van Elk, R., van Nierop, P., Smit, A., and Sixma, T. (2005) *J. Biol. Chem.* **280,** 26457–26466
7. Brejc, K., van Dijk, W. J., Klaassen, R. V., Schuurmans, M., van Der Oost, J., Smit, A. B., and Sixma, T. K. (2001) *Nature* **411,** 269–276
8. Celie, P., van Rossum-Fikkert, S., van Dijk, W., Brejc, K., Smit, A. B., and Sixma, T. (2004) *Neuron* **41,** 907–914
9. Hansen, S., Sulzenbacher, G., Huxford, T., Marchot, P., Taylor, P., and Bourne, Y. (2005) *EMBO J.* **24,** 3635–3646
10. Unwin, N. (2005) *J. Mol. Biol.* **346,** 967–989
11. Miyazawa, A., Fujiyoshi, Y., and Unwin, N. (2003) *Nature* **423,** 949–955
12. Colquhoun, D. (1998) *Br. J. Pharmacol.* **125,** 923–948
13. Chang, Y., and Weiss, D. (1999) *Biophys. J.* **77,** 2542–2551
14. Grosman, C., Zhou, M., and Auerbach, A. (2000) *Nature* **403,** 773–776
15. Absalom, N., Lewis, T., and Schofield, P. (2004) *Exp. Physiol.* **89,** 145–153
16. Lockless, S., and Ranganathan, R. (1999) *Science* **286,** 295–299
17. Hatley, M., Lockless, S., Gibson, S., Gilman, A., and Ranganathan, R. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 14445–14450
18. Suel, G., Lockless, S., Wall, M., and Ranganathan, R. (2003) *Nat. Struct. Biol.* **10,** 59–69
19. Shulman, A., Larson, C., Mangelsdorf, D., and Ranganathan, R. (2004) *Cell* **116,** 417–429
20. Gobel, U., Sander, C., Schneider, R., and Valencia, A. (1994) *Proteins* **18,** 309–317
21. Olmea, O., Rost, B., and Valencia, A. (1999) *J. Mol. Biol.* **293,** 1221
22. McLachlan, A. (1971) *J. Mol. Biol.* **61,** 409–424
23. Fodor, A., and Aldrich, R. (2004) *Proteins* **56,** 211–221
24. Olsen, R., Chang, C., Li, G., Hanchar, H., and Wallner, M. (2004) *Biochem. Pharmacol.* **68,** 1675–1684
25. Bera, A., Chatav, M., and Akabas, M. (2002) *J. Biol. Chem.* **277,** 43002–43010
26. Campos-Caro, A., Sala, S., Ballesta, J., Vicente-Agullo, F., Criado, M., and Sala, F. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6118–6123
27. Lynch, J., Rajendra, S., Pierce, K., Handford, C., Barry, P., and Schofield, P. (1997) *EMBO J.* **16,** 110–120
28. Rovira, J., Vicente-Agullo, F., Campos-Caro, A., Criado, M., Sala, F., Sala, S., and Ballesta, J. (1999) *Pfluegers Arch. Eur. J. Physiol.* **439,** 86–92
29. Kash, T., Jenkins, A., Kelley, J., Trodell, J. R., and Harrison, N. (2003) *Nature* **421,** 272–275
30. Kash, T., Dizon, M., Trudell, J., and Harrison, N. (2004) *J. Biol. Chem.* **279,** 4887–4893
31. Reeves, D., Jansen, M., Bali, M., Lemster, T., and Akabas, M. (2005) *J. Neurosci.* **25,** 9358–9366
32. Xiu, X., Hanek, A., Wang, J., Lester, H., and Dougherty, D. (2005) *J. Biol. Chem.* **280,** 41655–41666
33. Chakrapani, S., Bailey, T., and Auerbach, A. (2004) *J. Gen. Physiol.* **123,** 341–356
34. Bouzat, C., Gumilar, F., Spitzmaul, G., Wang, H., Rayes, D., Hansen, S., Taylor, P., and Sine, S. (2004) *Nature* **430,** 896–900
35. Lee, W., and Sine, S. (2005) *Nature* **438,** 243–247
36. Lummis, S., Beene, D., Lee, L., Lester, H., Broadhurst, R., and Dougherty, D. (2005) *Nature* **438,** 248–252
37. Boileau, A., and Czajkowski, C. (1999) *J. Neurosci.* **19,** 10213–10220
38. Davies, M., Newell, J., and Dunn, S. (2001) *J. Neurochem.* **79,** 55–62
39. Revah, F., Bertrand, D., Gaizi, J., Devillers-Thiery, A., Mulle, C., Hussy, N., Bertrand, S., Ballivet, M., and Changeux, J. (1991) *Nature* **353,** 846–849
40. Labarca, C., Nowak, M., Zhang, H., Tang, L., Deshpande, P., and Lester, H. (1995) *Nature* **376,** 514–516
41. Chang, Y., and Weiss, D. (1998) *Mol. Pharmacol.* **53,** 511–523
42. Galzi, J., Devillers-Thiery, A., Hussy, N., Bertrand, S., Changeux, J., and Bertrand, D. (1992) *Nature* **359,** 500–505
43. Gunthorpe, M., and Lummis, S. (2001) *J. Biol. Chem.* **276,** 10977–10983
44. Wotring, V., Miller, T., and Weiss, D. (2003) *J. Physiol.* **548,** 527–540
45. Keramides, A., Moorhouse, A., French, C., Schofield, P., and Barry, P. (2000) *Biophys. J.* **79,** 247–259
46. Jensen, M., Pedersen, L., Timmermann, D., Schousboe, A., and Ahring, P. (2005) *J. Neurochem.* **92,** 962–972
47. Pan, Z., Zhang, D., Zhang, X., and Lipton, S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6490–6495
48. Miko, A., Werby, E., Sun, H., Healey, J., and Zhang, L. (2004) *J. Biol. Chem.* **279,** 22833–22840
49. Breitinger, H., Lanig, H., Vohwinkel, C., Grewer, C., Breitinger, U., Clark, T., and

Becker, C. (2004) *Chem. Biol.* **11,** 1339–1350
50. Lo, D., Pinkham, J., and Stevens, C. (1991) *Neuron* **6,** 31–40
51. Lobitz, N., Gisselmann, G., Hatt, H., and Wetzel, C. (2001) *Mol. Pharmacol.* **59,** 844–851
52. De Rosa, M., Rayes, D., Spitzmaul, G., and Bouzat, C. (2002) *Mol. Pharmacol.* **62,** 406–414
53. Bouzat, C., Barrantes, F., and Sine, S. (2000) *J. Gen. Physiol.* **115,** 663–672
54. Bouzat, C., Gumilar, F., del Carmen Esandi, M., and Sine, S. (2002) *Biophys. J.* **82,** 1920–1929
55. Hemmings, H. C., Jr., Akabas, M., Goldstein, P., Trudell, J., Orser, B., and Harrison, N. (2005) *Trends Pharmacol. Sci.* **26,** 503–510
56. Chang, C., Olcese, R., and Olsen, R. (2003) *J. Biol. Chem.* **278,** 42821–42828
57. Mihic, S., Ye, Q., Wick, M., Koltchine, V., Krasowski, M., Finn, S., Mascia, M., Valenzuela, C., Hanson, K., Greenblatt, E., Harris, R., and Harrison, N. (1997) *Nature* **389,** 385–389
58. Nishikawa, K., and Harrison, N. (2003) *Anesthesiology* **99,** 678–684
59. Morris, K., and Amin, J. (2004) *Mol. Pharmacol.* **66,** 56–69
60. Amin, J. (1999) *Mol. Pharmacol.* **55,** 411–423
61. Pan, Z., Zhang, X., and Lipton, S. (2000) *Neuroscience* **98,** 333–338
62. Pratt, M., Husain, S., Miller, K., and Cohen, J. (2000) *J. Biol. Chem.* **275,** 29441–29451
63. Chang, Y., and Weiss, D. (1999) *Nat. Neurosci.* **2,** 219–225
64. Sedelnikova, A., Smith, C., Zakharkin, S., Davis, D., Weiss, D., and Chang, Y. (2005) *J. Biol. Chem.* **280,** 1535–1542
65. Boileau, A., Glen Newell, J., and Czajkowski, C. (2002) *J. Biol. Chem.* **277,** 2931–2937
66. Newell, J., McDevitt, R., and Czajkowski, C. (2004) *J. Neurosci.* **24,** 11226–11235
67. Torres, V., and Weiss, D. (2002) *J. Biol. Chem.* **277,** 43741–43748
68. Chang, Y., Wang, R., Barot, S., and Weiss, D. (1996) *J. Neurosci.* **16,** 5415–5424
69. Filizola, M., Olmea, O., and Weinstein, H. (2002) *Protein Eng.* **15,** 881–885

Evolutionarily Conserved Allosteric Network in the Cys Loop Family of Ligand-gated Ion Channels Revealed by Statistical Covariance Analyses

Yonghui Chen, Kevin Reilly and Yongchang Chang

J. Biol. Chem. 2006, 281:18184-18192.
doi: 10.1074/jbc.M600349200 originally published online April 4, 2006
