Molecular Neurobiology
© Copyright 1992 by The Humana Press Inc.
All rights of any nature whatsoever reserved.
ISSN 0893-7648/1992/4(3-4): 129–169/$2.00

Ligand-Gated Ion Channels

Homology and Diversity

V. B. Cockcroft,[1,2] D. J. Osguthorpe,[1] E. A. Barnard,[3] A. E. Friday,[4] and G. G. Lunt *,[2]

[1]Molecular Graphics Unit, Bath University, Bath BA2 7AY, UK; [2]Molecular Neuroscience Group, Department of Biochemistry, Bath University, Bath BA2 7AY, UK; [3]MRC Molecular Neurobiology Unit, MRC Center, Cambridge CB2 2QH, UK; [4]Zoology Department, University of Cambridge, Cambridge CB3 2JK, UK

Contents

*Author to whom all correspondence and reprint requests should be addressed.

Outline of the Ligand-Gated Ion Channel (LGIC) Superfamily

A current theme in biology is that of the gene superfamily. The concept of a superfamily encompasses a modal theme in which basic common features of the individual members of a superfamily may be grouped together, as indeed may features that are restricted to only a few members. In the case of neurotransmitter receptors, this overall unifying approach contrasts with the traditional approaches of pharmacology and physiology, which tend to emphasize the differences between receptor classes. The ligand-gated ion channels (LGICs) constitute such a superfamily, and in this review, it is considered to include the nicotinic acetylcholine (nACh) receptor, the glycine receptor and the $GABA_A$ receptor (*see* Karlin, 1980 and Barnard et al., 1987 for earlier reviews). Only in these three cases has homology at the level of primary structures been clearly established (Schofield et al., 1987; Grenningloh et al., 1987). These receptors share the common function of a regulated opening and closing of an ion channel integral to their structure, thereby facilitating, depending on the ion specificity of the channel, a depolarization or hyperpolarization of the membrane.

The first neurotransmitter receptor to be studied using molecular cloning was the nACh receptor of the electric organ of *Torpedo*, a cartilaginous fish. The gene sequence of the α, β-, γ-, and δ-subunits of this receptor was determined in 1982–83 (Noda et al., 1982,1983a; Claudio, 1983). This work revealed the full extent of the homology between each of the subunits of the *Torpedo* nACh receptor. The putative transmembrane segments were seen to be particularly well conserved, and using cDNA fragments over such regions, subunits of nACh receptors from the skeletal muscle of several vertebrates were isolated by crosshybridization (Noda et al., 1983b; Kubo et al., 1985; Nef et al., 1984; La Polla et al., 1984; Boulter et al., 1985). In addition, a novel bovine subunit was identified, termed ε, that could functionally substitute for the γ-subunit in expression studies (Takai et al., 1985). Cross-species comparisons of the subunit sequences showed the α-subunit more highly conserved than either the β-, γ-, or δ-subunits, and suggested that crosshybridization might also be used to clone nicotinic receptors and related proteins from neuronal tissues. Additionally, the α-subunit was seen to contain a unique pair of vicinal cysteines at positions 192–193. A neuronal subunit cloned from the PC12 cell line was termed α3 (Boulter et al., 1986), since it contained cysteines equivalent to those at position 192–193 of the α-subunit of the muscle nACh receptor. A second type of neuronal subunit was subsequently cloned from rat and termed β2, and was shown to be required along with α3 for the efficient expression of functional nACh receptors in the *Xenopus* oocyte system (Boulter et al., 1988). The chicken equivalent to β2 was termed nonα, to indicate that it lacked the characteristic cysteines equivalent to 192–193 of the muscle receptor α-subunit (Nef et al., 1988).

The α- and β-subunits of the $GABA_A$ receptor from bovine brain (Schofield et al., 1987) and the 48-kDa subunit of the glycine receptor from rat spinal cord (Grenningloh et al., 1987) were cloned in 1987. Analysis of these sequences showed them not only to be highly homologous to each other, but also significantly homologous to subunits of the nACh receptors. It was from these studies that the concept of the LGIC superfamily arose (Schofield et al., 1987; Grenningloh et al., 1987). Since these key events, there has been a rapid accumulation of sequence data for the superfamily. Sequences derived from subunit cDNA cloning are now known for muscle-type nACh receptors from fish electric organ and from skeletal muscle of mammals, birds, and amphibians, for neural nACh receptors from mammals, birds, and insects, and for $GABA_A$ receptors and glycine receptors from several mammals, giving >40 sequences in all, over a wide phylogenetic range (*see* Fig. 1, pp. 132–146).

Following the realization that the nACh receptor, $GABA_A$ receptor, and the glycine receptor

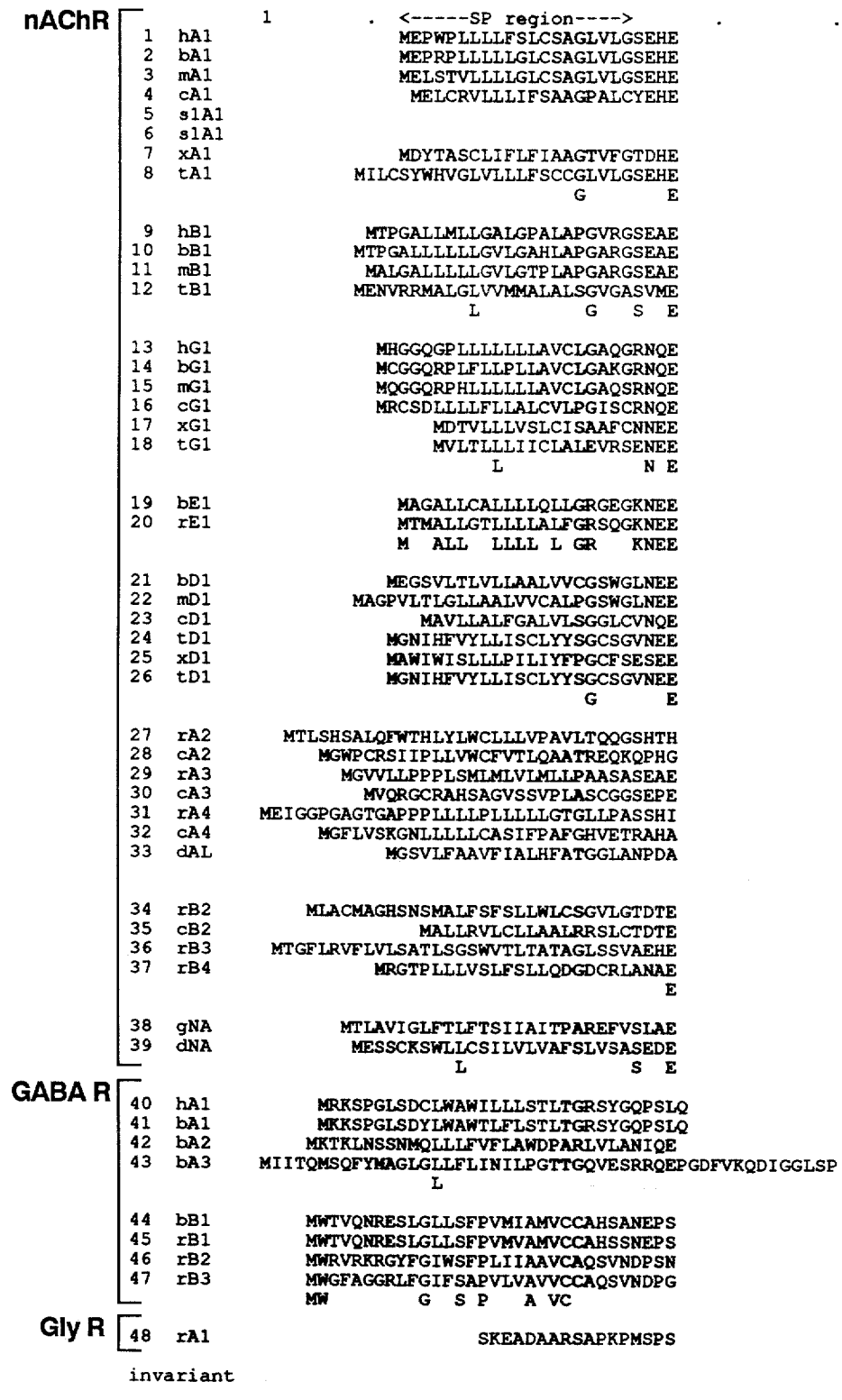
constitute a homologous group, it was considered possible that all ionotropic receptors would belong to this superfamily. However, with the recent cloning of a subunit of a kainate receptor from rat (Hollman et al., 1989), and of kainate binding proteins from chicken (Gregor et al., 1989) and frog (Wada et al., 1989), this assumption seems less secure. The kainate proteins are clearly related to each other, but it is less apparent that they belong to the above-defined LGIC superfamily. Despite similarities in the function and in the predicted transmembrane topologies of these two receptor types, none of the invariant residue positions in the kainate set of sequences occurs in the LGIC set, and vice-versa. For this reason, in this review these kainate binding proteins have not been considered as being part of the LGIC superfamily.

There are further difficulties with the definition of members of the ligand-gated receptor ion channel superfamily. Some new forms obtained by crosshybridization that are clearly related in sequence do not apparently function as activatable ion channels (Deneris et al., 1989; Bossy et al., 1988; Boulter et al., 1990). The oligomeric nature of LGICs also causes problems, particularly when several subunit subtypes are identified (Duvois et al., 1989; Levitan et al., 1988; Pritchett et al., 1989). In heterologous expression studies, functional receptors have been produced with just a single subunit subtype, as well as with combinations of different subunits (Pritchett et al., 1989; Blair et al., 1988; Schmieden et al., 1989). If this degree of flexibility in the formation of receptors, without fixed stoichiometries of subunits, occurs in vivo then complex mixtures of receptors may be formed. A further complication may be that, if each subunit brings to a hetero-oligomeric receptor its own binding site, mixed receptor forms could exist that are activated by more than one type of neurotransmitter. However, as yet no case is known of subunits for different transmitters coassembling in this way.

Other extended membrane protein superfamilies that are distinct from LGICs have been identified and deserve mention. These are:

1. The voltage-gated ion channels (Catterall, 1988), including the sodium (Noda et al., 1986), calcium (Tanabe et al., 1987), and potassium channels (Pongs et al., 1988)(and more recently the cGMP receptor ion channel [Kaupp et al., 1989; Jan and Jan, 1990])
2. The transport proteins (Higgins, 1989), such as the glucose transporter (Sarkar et al., 1988), the quinate transporter (Geever et al., 1989), the multidrug resistance proteins (Chen et al., 1986), the gene product of the cystic fibrosis locus (Riordan et al., 1989), and the GABA uptake transporter (Nelson et al., 1990)
3. The ryanodine (Takeshima et al., 1989) and the inositol trisphosphate receptors (Furuichi et al., 1989)
4. The second messenger linked receptors, which usually couple through G-proteins (Lefkowitz and Caron, 1988), examples of which are the multiple subtypes of muscarinic receptors (Chung and Venter, 1986), the catecholaminergic receptors (Bylund, 1988), and receptors for small peptides (Masu et al., 1987; Jackson et al., 1988), and
5. The growth-factor receptors (Carpenter and Zendegui, 1986).

Fig. 1. (following page) Alignment of LGIC subunit sequences. The amino acid sequences in Schofield et al., 1987; Grenningloh et al., 1987; Kubo et al., 1985; Nef et al., 1984,1988; LaPolla et al., 1984; Boulter et al., 1985; Takai et al., 1985; Deneris et al., 1989; Bossy et al., 1988; Levitan et al., 1988; Pritchett et al., 1989; Baldwin et al., 1988; Buonanno et al., 1989; Criado et al., 1988; Wada et al., 1988; Cauley et al., 1989; and Ymer et al., 1989 were taken from the OWL sequence database. Nomenclature: lowercase letter represents the species (h = human; r = rat; m = mouse; b = bovine; c = chicken; t = torpedo; d = drosophila). This is followed by an uppercase letter designating the subunit type (A = α; B = β; G = γ; D = δ; E = ε). The receptor types are shown in large lettering. A consensus sequence, showing conserved residues, is given under each group of sequences, and a grand consensus is shown at the bottom of the alignment. The position of the signal peptide (SP) region, the proposed calcium binding motif, the *cys*-loop, the cysteines 192–193, the transmem-brane segments M1–M4 and the Imoto ring positions (numbers 1–4 at the ends of M2) are highlighted above the alignment. The one-letter code for amino acids is used, and insertions are shown as gaps.


nAChR
GABA R
Gly R
1 . <-----SP region----> . .
1 hA1 MEPWPLLLLFSLCSAGLVLGSEHE
2 bA1 MEPRPLLLLLGLCSAGLVLGSEHE
3 mA1 MELSTVLLLLGLCSAGLVLGSEHE
4 cA1 MELCRVLLLIFSAAGPALCYEHE
5 s1A1
6 s1A1
7 xA1 MDYTASCLIFLFIAAGTVFGTDHE
8 tA1 MILCSYWHVGLVLLLFSCCGLVLGSEHE
G E
9 hB1 MTPGALLMLLGALGPALAPGVRGSEAE
10 bB1 MTPGALLLLLLGVLGAHLAPGARGSEAE
11 mB1 MALGALLLLLGVLGTPLAPGARGSEAE
12 tB1 MENVRRMALGLVVMMALALSGVGASVME
L G S E
13 hG1 MHGGQGPLLLLLLLAVCLGAQGRNQE
14 bG1 MCGGQRPLFLLPLLAVCLGAKGRNQE
15 mG1 MQGGQRPHLLLLLLAVCLGAQSRNQE
16 cG1 MRCSDLLLLFLLALCVLPGISCRNQE
17 xG1 MDTVLLLVSLCISAAFCNNEE
18 tG1 MVLTLLLIICLALEVRSENEE
L N E
19 bE1 MAGALLCALLLLQLLGRGEGKNEE
20 rE1 MTMALLGTLLLLALFGRSQGKNEE
M ALL LLLL L GR KNEE
21 bD1 MEGSVLTLVLLAALVVCGSWGLNEE
22 mD1 MAGPVLTLGLLAALVVCALPGSWGLNEE
23 cD1 MAVLLALFGALVLSGGLCVNQE
24 tD1 MGNIHFVYLLISCLYYSGCSGVNEE
25 xD1 MAWIWISLLLPILIYFPGCFSESEE
26 tD1 MGNIHFVYLLISCLYYSGCSGVNEE
G E
27 rA2 MTLSHSALQFWTHLYLWCLLLVPAVLTQQGSHTH
28 cA2 MGWPCRSIIPLLVWCFVTLQAATREQKQPHG
29 rA3 MGVVLLPPPLSMLMLVLMLLPAASASEAE
30 cA3 MVQRGCRAHSAGVSSVPLASCGGSEPE
31 rA4 MEIGGPGAGTGAPPPLLLLPLLLLLGTGLLPASSHI
32 cA4 MGFLVSKGNLLLLLCASIFPAFGHVETRAHA
33 dAL MGSVLFAAVFIALHFATGGLANPDA
34 rB2 MLACMAGHSNSMALFSFSLLWLCSGVLGTDTE
35 cB2 MALLRVLCLLAALRRSLCTDTE
36 rB3 MTGFLRVFLVLSATLSGSWVTLTATAGLSSVAEHE
37 rB4 MRGTPLLLVSLFSLLQDGDCRLANAE
E
38 gNA MTLAVIGLFTLFTSIIAITPAREFVSLAE
39 dNA MESSCKSWLLCSILVLVAFSLVSASEDE
L S E
40 hA1 MRKSPGLSDCLWAWILLLSTLTGRSYGQPSLQ
41 bA1 MKKSPGLSDYLWAWTLFLSTLTGRSYGQPSLQ
42 bA2 MKTKLNSSNMQLLLFVFLAWDPARLVLANIQE
43 bA3 MIITQMSQFYMAGLGLLFLINILPGTTGQVESRRQEPGDFVKQDIGGLSP
L
44 bB1 MWTVQNRESLGLLSFPVMIAMVCCAHSANEPS
45 rB1 MWTVQNRESLGLLSFPVMVAMVCCAHSSNEPS
46 rB2 MWRVRKRGYFGIWSFPLIIAAVCAQSVNDPSN
47 rB3 MWGFAGGRLFGIFSAPVLVAVVCCAQSVNDPG
MW G S P A VC
48 rA1 SKEADAARSAPKPMSPS
invariant

```
          51          .          .          .          .          .
 1  hA1                    TRLVAKLFKD  YSSVVRPVEDHRQVVEVTVGL
 2  bA1                    TRLVAKLFED  YNSVVRPVEDHRQAVEVTVGL
 3  mA1                    TRLVAKLFED  YSSVVRPVEDHREIVQVTVGL
 4  cA1                    TRLVDDLFRE  YSKVVRPVENHRDAVVVTVGL
 5  s1A1
 6  s1A1
 7  xA1                    TRLIGDLFAN  YNKVVRPVETYKDQVVVTVGL
 8  tA1                    TRLVANLLEN  YNKVIRPVEHHTHFVDITVGL
                           TRLV  L     Y  V RPVE     V  TVGL

 9  hB1                    GRLREKLFSG  YDSSVRPAREVGDRVRVSVGL
10  bB1                    GRLREKLFSG  YDSTVRPAREVGDRVWVSIGL
11  mB1                    GQLIKKLFSN  YDSSVRPAREVGDRVGVSIGL
12  tB1                    DTLLSVLFET  YNPKVRPAQTVGDKVTVRVGL
                             L   LF    Y   VRPA  VGDKV V  GL

13  hG1                    ERLLADLMQN  YDPNLRPAERDSDVVNVSLKL
14  bG1                    ERLLGDLMQG  YNPHLRPAEHDSDVVNVSLKL
15  mG1                    ERLLADLMRN  YDPHLRPAERDSDVVNVSLKL
16  cG1                    EKLLQDLMTN  YNRHLRPALRGDQVIDVTLKL
17  xG1                    ERLLNDLMKN  YNKNLRPVEKDGDIISVSIKL
18  tG1                    GRLIEKLLGD  YDKRIIPAKTLDHIIDVTLKL
                            RL   LL    Y     P          V  KL

19  bE1                    LRLYHYLFDT  YDPGRRPVQEPEDTVTISLKV
20  rE1                    LSLYHHLFDN  YDPECRPVRRPEDTVTITLKV
                           L LYH LFD   YDP  RPV  PEDTVTI LKV

21  bD1                    ERLIRHLFEEKAYNKELRPAAH KESVEISLAL
22  mD1                    QRLIQHLFNEKGYDKDLRPVARKEDKVDVALSL
23  cD1                    ERLIHHLFEERGYNKEVRPVASADEVVDVYLAL
24  tD1                    ERLINDLLIVNKYNKHVRPVKHNNEVVNIALSL
25  xD1                    ERLLNHIFVERGYRKELRPVEHTGETVNVSLAL
26  tD1                    ERLINDLLIVNKYNKHVRPVKHNNEVVNIALSL
                            RLI        Y K  RP       V   L L

27  rA2                 AEDRLFKHLFGG  YNRWARPVPNTSDVVIVRFGL
28  cA2                FAEDRLFKHLFTG  YNRWSRPVPNTSDVVIVKFGL
29  rA3                   HRLFQYLFED  YNEIIRPVANVSHPVIIQFEV
30  cA3                   HRLYAALFKN  YNQFVRPVKNASDPVIIQFEV
31  rA4             ETRAHAEERLLKRLFSG  YNKWSRPVGNISDVVLVRFGL
32  cA4                EERLLKKLFSG  YNKWSRPVANISDVVLVRFGL
33  dAL                   KRLYDDLLSN  YNRLIRPVGNNSDRLTVKMGL
                           RL   L     YN   RPV N S

34  rB2                    ERLVEHLLDPSRYNKLIRPATNGSELVTVQLMV
35  cB2                    ERLVEYLLDPTRYNKLIRPATNGSQLVTVQLMV
36  rB3                    DALLRHLFQG  YQKWVRPVLNSSDIIKVYFGL
37  rB4                    EKLMDDLLNKTRYNNLIRPATSSSQLISIRLEL
                             L   L     Y    RP    S

38  gNA                 REDALLRELFQG  YQRWVRPVQHANHSVKVRFGL
39  dNA                   ERLVRDLFRG  YNKLIRPVQNMTQKVGVRFGL
                            L R LF G  Y    RPVQ     V VRFGL

40  hA1           DELKDNTTVFTRILDRLLDG  YDNRLRPGL GERVTEVKTDI
41  bA1           DELKDNTTVFTRILDRLLDG  YDNRLRPGL GERVTEVKTDI
42  bA2           DEAKNNITIFTRILDRLLDG  YDNRLRPGL GDSITEVFTNI
43  bA3    KHAPDIPDDSTDNITIFTRILDRLLDG  YDNRLRPGL GDAVTEVKTDI
                 D    NIT FTRILDRLLDG  YDNRLRPGL G   TEVKT I

44  bB1               NMSYVKETVDRLLKG  YDIRLRP DFGGPPVDVGMRI
45  rB1               NMSYVKETVDRLLKG  YDIRLRP DFGGPPVDVGMRI
46  rB2                MSLVKETVDRLLKG  YDIRLRP DFGGPPVAVGMNI
47  rB3               NMSFVKETVDKLLKG  YDIRLRP DFGGPPVCVGMNI
                       MS VKETVD LLKG  YDIRLRP DFGGPPV VGM I

48  rA1                    DFLDKLMGRTSGYDARIRP NFKGPPVNVSCNI

invariant                              Y     P
```

```
          101
                .         .         .         .         .
 1  hA1    QLIQLINVDEVNQIVTTNVRLKQQWVDYNLKWNPDDYGGVKKIHIPSEKI
 2  bA1    QLIQLINVDEVNQIVTTNVRLKQQWVDYNLKWNPDDYGGVKKIHIPSEKI
 3  mA1    QLIQLINVDEVNQIVTTNVRLKQQWVDYNLKWNPDDYGGVKKIHIPSEKI
 4  cA1    QLIQLINVDEVNQIVTTNVRLKQQWTDINLKWNPDDYGGVKQIRIPSDDI
 5  s1A1
 6  s1A1
 7  xA1    QLIQLINVDEVNQIVSTNIRLKQQWRDVNLKWDPAKYGGVKKIRIPSSDV
 8  tA1    QLIQLISVDEVNQIVETNVRLRQQWIDVRLRWNPADYGGIKKIRLPSDDV
           QLIQLI VDEVNQIV TN RL QQW D  L W P  YGG KKI  PS

 9  hB1    ILAQLISLNEKDEEMSTKVYLDLEWTDYRLSWDPAEHEGIDSLRITAESV
10  bB1    TLAQLISLNEKDEEMSTKVYLDLEWTDYRLSWDPEEHEGIDSLRISAESV
11  mB1    TLAQLISLNEKDEEMSTKVYLDLEWTDYRLSWDPAEHDGIDSLRITAESV
12  tB1    TLTNLLILNEKIEEMTTNVFLNLAWTDYRLQWDPAAYEGIKDLRIPSSDV
            L  L  LNEK EEM T V L   WTDYRL WDP    GI  LRI    V

13  hG1    TLTNLISLNEREEALTTNVWIEMQWCDYRLRWDPRDYEGLWVLRVPSTMV
14  bG1    TLTNLISLNEREEALTTNVWIEMQWCDYRLRWDPRDYGGLWVLRVPSTMV
15  mG1    TLTNLISLNEREEALTTNVWIEMQWCDYRLRWDPKDYEGLWILRVPSTMV
16  cG1    TLTNLISLNEREETLTTNVWIEMQWSDYRLRWDPDKYDDIQQLRVPSAMV
17  xG1    TLTNLISLNEKEEALTTNVWVEMQWKDYRLSWDPNDYHGISMMRIPSTSV
18  tG1    TLTNLISLNEKEEALTTNVWIEIQWNDYRLSWNTSEYEGIDLVRIPSELL
           TLTNLISLNEKEE LTTNVWIE QW DYRL W    Y      RIPS  L

19  bE1    TLTNLISLNEKEETLTTSVWIGIDWQDYRLNYSKGDFGGVETLRVPSELV
20  rE1    TLTNLISLNEKEETLTTSVWIGIEWQDYRLNFSKDDFAGVEILRVPSEHV
           TLTNLISLNEKEETLTTSVWIGI WQDYRLN SK DF GVEILRVPSE V

21  bD1    TLSNLISLKEVEETLTTNVWIEQGWTDSRLQWDAEDFGNISVLRLPADMV
22  mD1    TLSNLISLKEVEETLTTNVWIDHAWVDSRLQWDANDFGNITVLRLPPDMV
23  cD1    TLSNLISLKEVDETLTTNVWVEQSWTDYRLQWNTSEFGGVDVLRLLPEML
24  tD1    TLSNLISLKETDETLTSNVWMDHAWYDHRLTWNASEYSDISILRLPPELV
25  xD1    TLSNLISLKEADETLTTNVWVELAWYDKRLAWDMETYNNIDILRVPPDMV
26  tD1    TLSNLISLKETDETLTSNVWMDHAWYDHRLTWNASEYSDISILRLPPELV
           TLSNLISLKE DETLT NVW    W D RL W          LR  P

27  rA2    SIAQLIDVDEKNQMMTTNVWLKQEWNDYNVRWDPAEFGNVTSLRVPSEMI
28  cA2    SIAQLIDVDEKNQMMTTNVWLKQEWSDYKLRWNPEDFDNVTSIRVPSEMI
29  rA3    SMSQLVKVDEVNQIMETNLWLKQIWNDYKLKWKPSDYQGVEFMRVPAEKI
30  cA3    SMSQLVKVDEVNQIMETNLWLKHIWNDYKLRWNPVDYGGAEFIRVPSGQI
31  rA4    SIAQLIDVDEKNQMMTTNVWVKQEWHDYKLRWDPGDYENVTSIRIPSELI
32  cA4    SIAQLIDVDEKNQMMTTNVWVKQEWHDYKLRWDPQEYENVTSIRIPSELI
33  dAL    RLSQLIDVNLKNQIMTTNVWVEQEWNDYKLKWNPDDYGGVDTLHVPSEHI
              QL  V  KNQ M TN W  Q W DYK  W P           P   I

34  rB2    SLAQLISVHEREQIMTTNVWLTQEWEDYRLTWKPEDFDNMKKVRLPSKHI
35  cB2    SLAQLISVHEREQIMTTNVWLTQEWEDYRLTWKPEDFDNMKKVRLPSKHI
36  rB3    KISQLVDVDEKNQLMTTNVWLKQEWTDQKLRWNPEEYGGINSIKVPSESL
37  rB4    SLSQLISVNEREQIMTTSIWLKQEWTDYRLAWNSSCYEGVNILRIPARRV
              QL  V E  Q MTT  WL QEW D  L W             P

38  gNA    KISQLVDVDEKNQLMTTNVWLWQEWLDYKLRWNPENYGGITSIRVPSESI
39  dNA    AFVQLINVNEKNQVMKSNVWLRLVWYDYQLQWDEADYGGIGVLRLPPDKV
              QL  V EKNQ M  NVWL   W DY L W    YGGI   R P

40  hA1    FVTSFGPVSDHDMEYTIDVFFRQSWKDERLKFK GPMTVLRLNNLMASKI
41  bA1    FVTSFGPVSDHDMEYTIDVFFRQSWKDERLKFK GPMTVLRLNNLMASKI
42  bA2    YVTSFGPVSDTDMEYTIDVFFRQKWKDERLKFK GPMNILRLNNLMASKI
43  bA3    YVTSFGPVSDTDMEYTIDVFFRQTWHDERLKFD GPMKILPLNNLLASKI
           VTSFGPVS  DMEYTIDVFFRQ W DERLKF  GPM  LPLNNL ASKI

44  bB1    DVASIDMVSEVNMDYTLTMYFQQSWKDKRLSYSGIPLNLTLDNRV ADQL
45  rB1    DVASIDMVSEVNMDYTLTMYFQQSWKDKRLSYSGIPLNLTLDNRV ADQL
46  rB2    DIASIDMVSEVNMDYTLTMYFQQAWRDKRLSYWVIPLNLTLDNRV ADQL
47  rB3    DIASIDMVSEVNMDYTLTMYFQQYWRDKRLAYSGIPLNLTLDNRV ADQL
           D ASIDMVSEVNMDYTLTMYFQQ W DKRL Y  IPLNLTLDNRV ADQL

48  rA1    FINSFGSIAETTMDYRVNIFLRQQWNDPRLAYMEYPDDSLDLDPSMLDSI

invariant                          W D  L
```

```
          151      // Ca 2+ ///.          .          .    C_____
 1   hA1    WRPDLVLYNNADGDFAIVKFTKVLLQ  YTGHITWTPPAIFKSYCEIIVT
 2   bA1    WRPDLVLYNNADGDFAIVKFTKVLLD  YTGHITWTPPAIFKSYCEIIVT
 3   mA1    WRPDVVLYNNADGDFAIVKFTKVLLD  YTGHITWTPPAIFKSYCEIIVT
 4   cA1    WRPDLVLYNNADGDFAIVKYTKVLLE  HTGKITWTPPAIFKSYCEIIVT
 5   s1A1                               NPPAIFKSYCEIIVT
 6   s1A1                               NPPAIFKSYCEIIVT
 7   xA1    WSPDLVLYNNADGDFAISKDTKILLE  YTGKITWTPPAIFKSYCEIIVT
 8   tA1    WLPDLVLYNNADGDFAIVHMTKLLLD  YTGKIMWTPPAIFKSYCEIIVT
            W PD VLYNNADGDFAI   TK LL    TGKI W PPAIFKSYCEIIVT

 9   hB1    WLPDVVLLNNNDGNFDVALDISVVVS  SDGSVRWQPPGIYRSSCSIQVT
10   bB1    WLPDVVLLNNNDGNFDVALDINVVVS  SDGSMRWQPPGIYRSSCSIQVT
11   mB1    WLPDVVLLNNNDGNFDVALDINVVVS  FEGSVRWQPPGLYRSSCSIQVT
12   tB1    WQPDIVLMNNNDGSFEITLHVNVLVQ  HTGAVSWQPSAIYRSSCTIKVM
            W PD VL NNNDG F   L   V V     G   WQP    YRSSC I V

13   hG1    WRPDIVLENNVDGVFEVALYCNVLVS  PDGCIYWLPPAIFRSACSISVT
14   bG1    WRPDIVLENNVDGVFEVALYCNVLVS  PDGCVYWLPPAIFRSSCPVSVT
15   mG1    WRPDIVLENNVDGVFEVALYCNVLVS  PDGCIYWLPPAIFRSSCSISVT
16   cG1    WLPDIVLENNIDGTFEITLYTNVLVY  PDGSIYWLPPAIYRSSCSIHVT
17   xG1    WLPDVGLENNVDGTFDIALYTNTLVS  SDGSMYWLPPAIYRSSCPVVVT
18   tG1    WLPDVVLENNVDGQFEVAYYANVLVY  NDGSMYWLPPAIYRSTCPIAVT
            W PD  LENN DG F    Y N LV    DG  YWLPPAI RS C   VT

19   bE1    WLPEIVLENNIDGQFGVAYEANVLVS  EGGYLSWLPPAIYRSTCAVEVT
20   rE1    WLPEIVLENNIDGQFGVAYDCNVLVY  EGGSVSWLPPAIYRSTCAVEVT
            WLPEIVLENNIDGQFGVAY  NVLV   EGG  SWLPPAIYRSTCAVEVT

21   bD1    WLPEIVLENNNDGSFQISYSCNVLIY  PSGSVYWLPPAIFRSSCPISVT
22   mD1    WLPEIVLENNNDGSFQISYACNVLVY  DSGYVTWLPPAIFRSSCPISVT
23   cD1    WLPEIVLENNNDGLFEVAYYCNVLVY  NTGYVYWLPPAIFRSACPINVN
24   tD1    WIPDIVLQNNNDGQYHVAYFCNVLVR  PNGYVTWLPPAIFRSSCPINVL
25   xD1    WQPQLILENNNNGVFEVAYYSNVLIS  SDGFMYWLPPAIFQTSCSINVN
26   tD1    WIPDIVLQNNNDGQYHVAYFCNVLVR  PNGYVTWLPPAIFRSSCPINVL
            W P   L NNN G     Y  NVL     G   WLPPAIF   C I V

27   rA2    WIPDIVLYNNADGEFAVTHMTKAHLF  FTGTVHWVPPAIYKSSCSIDVT
28   cA2    WIPDIVLYNNADGEFAVTHMTKAHLF  SNGKVKWVPPAIYKSSCSIDVT
29   rA3    WKPDIVLYNNADGDFQVDDKTKALLK  YTGEVTWIPPAIFKSSCKIDVT
30   cA3    WKPDIVLYNNAVGDFQVDDKTKALLK  YTGDVTWIPPAIFKSSCKIDVT
31   rA4    WRPDIVLYNNADGDFAVTHLTKAHLF  YDGRVQWTPPAIYKSSCSIDVT
32   cA4    WRPDIVLYNNADGDFAVTHLTKAHLF  YDGRIKWMPPAIYKSSCSIDVT
33   dAL    WHPDIVLYNNADGNYEVTIMTKAILH  HTGKVVWKPPAIYKSFCEIDVE
            W PDIVLYNNA G   V   TKA L     G   W PPAI KSFC IDV

34   rB2    WLPDVVLYNNADGMYEVSFYSNAVVS  YDGSIFWLPPAIYKSACKIEVK
35   cB2    WLPDVVLYNNADGMYEVSFYSNAVIS  YDGSIFWLPPAIYKSACKIEVK
36   rB3    WLPDIVLFENADGRFEGSLMTKAIVK  SSGTVSWTPPASYKSSCTMDVT
37   rB4    WLPDIVLYNNADGTYEVSVYTNVIVR  SNGSIQWLPPAIYKSACKIEVK
            WLPD VL  NADG  E S            G   W PPA YKS C   V

38   gNA    WLPDIVLYENADGRFEGSLMTKAIVR  YNGMITWTPPASYKSACTMDVT
39   dNA    WKPDIVLFNNADGNYEVRYKSNVLIY  PTGEVLWVPPAIYQSSCTIDVT
            W PDIVL  NADG  E              G   W PPA Y S C  DVT

40   hA1    WTPDTFM              PNKLLRITEDGTLLYTMRLTVRAECPMHLE
41   bA1    WTPDTFFHNGKKSVAHNMTMPNKLLRITEDGTLLYTMRLTVRAECPMHLE
42   bA2    WTPDTFFHNGKKSVAHNMTMPNKLLRIQDDGTLLYTMRLTVQAECPMHLE
43   bA3    WTPDTFFHNGKKSVAHNMTTPNKLLRLVDNGTLLYTMRLTIHAECPMHLE
            WTPDTF HNGKKSVAHNMT PNKLLR    GTLLYTMRLT  AECPMHLE

44   bB1    WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLR
45   rB1    WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLR
46   rB2    WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLR
47   rB3    WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLR
            WVPDTYFLNDKKSFVHGVTVKNRMIRLHPDGTVLYGLRITTTAACMMDLR

48   rA1    WKPDLFFANEKGAHFHEITTDNKLLRISRNGNVLYSIRITLTLACPMDLK

invariant  W P                          G            C
```

```
          201_____C.          .          .          .
 1  hA1    HFPFDEQNCSMKLGTWTYDGSVVAINP                ESDQPDLSN
 2  bA1    HFPFDEQNCSMKLGTWTYDGSVVVINP                ESDQPDLSN
 3  mA1    HFPFDEQNCSMKLGTWTYDGSVVAINP                ESDQPDLSN
 4  cA1    YFPFDQQNCSMKLGTWTYDGTMVVINP                ESDRPDLSN
 5  s1A1   YFPFDEQNCSMKLGTRTYDGTVVAIYP                EGPRPDLSN
 6  s1A1   YFPFDEQNCSMKLGTWTYDGTVVAIYP                EGPRPDLSN
 7  xA1    YFPFDQQNCSMKFGTWTYDGSLLVINP                ERDRPDLSN
 8  tA1    HFPFDQQNCTMKLGIWTYDGTKVSISP                ESDRPDLST
           FPFD QNC MKLG  TYDG    I P                 E   PDLS

 9  hB1    YFPFDWQNCTMVFSSYSYDSSEVTLQTGLGPDGQG      TQEIHIHEGT
10  bB1    YFPFDWQNCTMVFSSYSYDSSEVSLQTGLSPEGQE      RQEVYIHEGT
11  mB1    YFPFDWQNCTMVFSSYSYDSSEVSLKTGLDPEGEE      RQEVYIHEGT
12  tB1    YFPFDWQNCTMVFKSYTYDTSEVTLQHALDAKGERE     VKEIVINKDA
           YFPFDWQNCTMVF SY YD SEV L   L   G          E  I

13  hG1    YFPFDWQNCSLIFQSQTYSTNEIDLQLSQEDGQT       IEWIFIDPEA
14  bG1    FFPFDWQNCSLIFQSQTYSTNEINLQLSQEDGQT       IEWIFIDPEA
15  mG1    YFPFDWQNCSLIFQSQTYSTSEINLQLSQEDGQA       IEWIFIDPEA
16  cG1    YFPFDWQNCTMVFQSQTYSANEINLLLTVEEGQT       IEWIFIDPEA
17  xG1    YFPFDWQNCSIVFQSQTYSANEIELLLTVDEQT        IEWIEIDPEA
18  tG1    YFPFDWQNCSLVFRSQTYNAHEVNLQLSAEEGEA       VEWIHIDPED
           FPFDWQNC   F SQTY   E  L L                EWI IDPE

19  bE1    YFPFDWQNCSLVFRSQTYNAEEVEFVFAVDDEGKT      ISKIDIDTEA
20  rE1    YFPFDWQNCSLIFRSQTYNAEEVELIFAVDDDGNA      INKIDIDTAA
           YFPFDWQNCSL FRSQTYNAEEVE  FAVDD G        I KIDIDT A

21  bD1    YFPFDWQNCSLKFSSLKYTTKEITLSLKQAEEDGRS  YPVEWIIIDPEG
22  mD1    YFPFDWQNCSLKFSSLKYTAKEITLSLKQEEENNRS  YPIEWIIIDPEG
23  cD1    FFPFDWQNCTLKFSSLAYNAQEINMHLKEESDPETEKNYRVEWIIIDPEG
24  tD1    YFPFDWQNCSLKFTALNYDANEITMDLMTDTIDGKD  YPIEWIIIDPEA
25  xD1    YFPFDWQNCSLKFSSLTYNAKEINLQLRQDLDEASQRYYPVEWIIIDPEG
26  tD1    YFPFDWQNCSLKFTALNYDANEITMDLMTDTIDGKD  YPIEWIIIDPEA
     1      FPFDWQNC LKF  L Y   EI   L            YP EWIIIDPE

27  rA2    FFPFDQQNCKMKFGSWTYDKAKIDLEQ               MERTVDLKD
28  cA2    YFPFDQQNCKMKFGSWTYDKAKIDLEN               MEHHVDLKD
29  rA3    YFPFDYQNCTMKFGSWSYDKAKIDLVL               IGSSMNLKD
30  cA3    YFPFDYQNCTMKFGSWSYDKAKIDLVL               IGSTMNLKD
31  rA4    FFPFDQQNCTMKFGSWTYDKAKIDLVS               IHSRVDQLD
32  cA4    FFPFDQQNCKMKFGSWTYDKAKIDLVS               MHSHVDQLD
33  dAL    YFPFDEQTCFMKFGSWTYDGYMVDLRHLKQTADSD       NIEVGIDLQD
           FPF  Q C MKFGSW YD    DL                        D

34  rB2    HFPFDQQNCTMKFRSWTYDRTEIDLVL               KSDVASLDD
35  cB2    HFPFDQQNCTMKFRSWTYDRTEIDLVL               KSEVASLDD
36  rB3    FFPFDRQNCSMKFGSWTYDGTMVDLIL               INENVDRKD
37  rB4    HFPFDQQNCTLKFRSWTYDHTEIDMVL               KSPTAIMDD
           FPFD QNC  KF SWTYD T  D  L                     D

38  gNA    FFPFDRQNCSMKFGSWTYDGNMVKLVL               INQQVDRSD
39  dNA    YFPFDQQTCIMKFGSWTFNGDQVSLALYN             NKNFVDLSD
           FPFD Q C MKFGSWT  G  V L L                   VD  D

40  hA1    DFPMDAHACPLKFGSYAYTRAEVVYEW                TREPAHSV
41  bA1    DFPMDAHACPLKFGSYAYTRAEVVYEW                TREPARSV
42  bA2    DFPMDAHSCPLKFGSYAYTTSEVTYIW                TYNASDSV
43  bA3    DFPMDVHACPLKFGSYAYTTAEVVYSW                TLGKNKSV
           DFPMD H CPLKFGSYAYT  EV Y W                T    SV

44  bB1    RYPLDEQNCTLEIESYGYTTDDIEFYW                  NGGEGA
45  rB1    RYPLDEQNCTLEIESYGYTTDDIEFYW                  NGGEGA
46  rB2    RYPLDEQNCTLEIESYGYTTDDIEFYW                  RGDDNA
47  rB3    RYPLDEQNCTLEIESYGYTTDDIEFYW                  RGGDKA
           RYPLDEQNCTLEIESYGYTTDDIEFYW                   G   A

48  rA1    NFPMDVQTCIMQLESFGYTMNDLIFEWQ               EQGAVQVAD

invariant    P D   C
```

```
          251        .          .<192-193>.          .    <------
 1  hA1    FMESGEWVIKESRGWKHSVT  YSCCPDTPYLDITYHFVMQRLPLYFIVN
 2  bA1    FMESGEWVIKESRGWKHWVF  YACCPSTPYLDITYHFVMQRLPLYFIVN
 3  mA1    FMESGEWVIKEARGWKHWVF  YSCCPTTPYLDITYHFVMQRLPLYFIVN
 4  cA1    FMESGEWVMKDYRGWKHWVY  YACCPDTPYLDITYHFLMQRLPLYFIVN
 5  s1A1   YMQSGEWALKDYRGFWHSVN  YSCCLDTPYLDITYHFILLRLPLYFIVN
 6  s1A1   YMQSGEWTLKDYRGFWHSVN  YSCCLDTPYLDITYHFILLRLPLYFIVN
 7  xA1    FMASGEWMMKDYRCWKHWVY  YTCCPDKPYLDITYHFVLQRLPLYFIVN
 8  tA1    FMESGEWVMKDYRGWKHWVY  YTCCPDTPYLDITYHFIMQRIPLYFVVN
            M SGEW  KD RG  H V   Y CC   PYLDITYHF   R PLYFV N

 9  hB1    FIENGQWENIHKPSRLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFYLVN
10  bB1    FIENGQWEIIHKPSRLIQPSVDPRGGGEGRREEVTFYLIIRRKPLFYLVN
11  mB1    FIENGQWEIIHKPSRLIQLPGDQRGGKEGHHEEVIFYLIIRRKPLFYLVN
12  tB1    FTENGQWSIEHKPSRKNWRS      DDPSYEDVTFYLIIQRKPLFYIVY
           FTENGQW   HKPSR                 VTFYLII RKPLFY V

13  hG1    FTENGEWAIQHRPAKMLLDPA  APAQEAGHQKVVFYLLIQRKPLFYVIN
14  bG1    FTENGEWAIRHRPAKMLLDEA  APAEEAGHQKVVFYLLIQRKPLFYVIN
15  mG1    FTENGEWAIRHRPAKMLLDSV  APAEEAGHQKVVFYLLIQRKPLFYVIN
16  cG1    FTENGEWAIKHRPARKIINSG RFTPDDIQYQQVIFYLIIQRKPLFYIIN
17  xG1    FTENGEWAIKHMPAKRIINHR  LPRDDVNYQQIVFYLIIQRKPLFYIIN
18  tG1    FTENGEWTIRHRPAKKNYNWQ  LTKDDTDFQEIIFFLIIQRKPLFYIIN
           FTENGEW I H PA                 Q   F LIIQRKPLFYIIN

19  bE1    YTENGEWAIDFCPGVIRRHDG  DSAGGPGETDVIYSLIIRRKPLFYVIN
20  rE1    FTENGEWAIDYCPGMIRHYEG  GSTEDPGETDVIYTLIIRRKPLFYVIN
            TENGEWAID CPG IR   G   S   PGETDVIY LIIRRKPLFYVIN

21  bD1    FTENGEWEIVHRPARVNVDPS  VPLDSPNRQDVTFYLIIRRKPLFYVIN
22  mD1    FTENGEWEIVHRAAKLNVDPS  VPMDSTNHQDVTFYLIIRRKPLFYIIN
23  cD1    FTENGEWEIIHRPARKNIHPS  YPTESSEHQDITFYLIIKRKPLFYVIN
24  tD1    FTENGEWEIIHKPAKKNIYPD  KFPNGTNYQDVTFYLIIRRKPLFYVIN
25  xD1    FTENGEWEIVHIPAKKNIDRS  LSPESTKYQDITFYLIIERKPLFYIIN
26  tD1    FTENGEWEIIHKPAKKNIYPD  KFPNGTNYQDVTFYLIIRRKPLFYVIN
           FTENGEWEI H  A  N              QD TFYLII  KPLFY IN

27  rA2    YWESGEWAIINATGTYNSKK   YDCCAEIYPDVTYYFVIRRLPLFYTIN
28  cA2    YWESGEWAIINAIGRYNSKK   YDCCTEIYPDITFYFVIRRLPLFYTIN
29  rA3    YWESGEWAIIKAPGYKHEIK   YNCCEEIYQDITYSLYIRRLPLFYTIN
30  cA3    YWESGEWAIIKAPGYKHDIK   YNCCEEIYTDITYSLYIRRLPLFYTIN
31  rA4    FWESGEWVIVDAVGTYNTRK   YECCAEIYPDITYAFIIRRLPLFYTIN
32  cA4    YWESGEWVIINAVGNYNSKK   YECCTEIYPDITYSFIIRRLPLFYTIN
33  dAL    YYISVEWDIMRVPAVRNEKF   YSCCEEPYLDIVFNLTLRRKTLFYTVN
              S EW I                Y CC E Y DI      RR  LFYT N

34  rB2    FTPSGEWDIIALPGRRNENP     DDSTYVDITYDFIIRRKPLFYTIN
35  cB2    FTPSGEWDIVALPGRRNENP     DDSTYVDITYDFIIRRKPLFYTIN
36  rB3    FFDNGEWEILNAKGMKGNRR     EGFYSYPFVTYSFVLRRLPLFYTLF
37  rB4    FTPSGEWDIVALPGRRTVNP     QDPSYVDVTYDFIIKRKPLFYTIN
           F   GEW I    G     P         Y   TY F   R PLFYT

38  gNA    FFDNGEWEILSATGVKGSRQ     DSHLSYPYITYSFILKRLPLFYTLF
39  dNA    YWKSGTWDIIEVPAYLNVYE   GDSNHPTETDITFYIIIRRKTLFYTVN
              G W I     Y            H     I    I  R  LFYT

40  hA1    VVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHFHLKRKIGYFVIQ
41  bA1    VVAEDGSRLNQYDLLGQTVDSGIVQSSTGEYVVMTTHFHLKRKIGYFVIQ
42  bA2    QVAPDGSRLNQYDLPGQSIGKETIKSSTGEYTVMTAHFHLKRKIGYFVIQ
43  bA3    EVAQDGSRLNQYDLLGHVVGTEIIRSSTGEYVVMTTHFHLKRKIGYFVIQ
           VA DGSRLNQYDL G          SSTGEY VMT HFHLKRKIGYFVIQ

44  bB1    VTGVNKIELPQFSIVDYKMVSKKVEFTTGAYPRLSLSFRLKRNIGYFILQ
45  rB1    VTGVNKIELPQFSIVDYKMVSKKVEFTTGAYPRLSLSFRLKRNIGYFILQ
46  rB2    VTGVTKIELPQFSIVDYKLITKKVVFSTGSYPRLSLSFKLKRNIGYFILQ
47  rB3    VTGVERIELPQFSIVEHRLVSRNVVFATGAYPRLSLSFRLKRNIGYFILQ
           VTGV  IELPQFSIV        V F TG YPRLSLSF LKRNIGYFILQ

48  rA1    GLTLPQFILKEEKDLRYCTK    HYNTGKFTCIEARFHLERQMGYYLIQ

invariant                                          R
```

```
                                     1   2                      3   4
              301----M1-----------> |   |-----------M2-------|-->.|
 1  hA1         VIIPCLLFSFLTGLVFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS
 2  bA1         VIIPCLLFSFLTGLVFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS
 3  mA1         VIIPCLLFSFLTSLVFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS
 4  cA1         VIIPCLLFSFLTGFVFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS
 5  slA1        VIIPC
 6  slA1        VIIPC
 7  xA1         VIIPCLLFSFLTGLVFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS
 8  tA1         VIIPCLLFSFLTGLVFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS
                VIIPCLLFSFLT  VFYLPTDSG EKMTLSISVLLSLTVFLLVIVELIPS

 9  hB1         VIAPCILITLLAIFVFYLPPDAG EKMGLSIFALLTLTVFLLLLADKVPE
10  bB1         VIAPCILITLLAIFVFYLPPDAG EKMGLSIFALLTLTVFLLLLADKVPE
11  mB1         VIAPCILITLLAIFVFYLPPDAG EKMGLSIFALLTLTVFLLLLADKVPE
12  tB1         TIIPCILISILAILVFYLPPDAG EKMSLSISALLAVTVFLLLLADKVPE
                 I PCILI  LAILVFYLPPDAG EKM LSI ALL  TVFLLLLADKVPE

13  hG1         IIAPCVLISSVAILIHFLPAKAGGQKCTVAINVLLAQTVFLFLLAKKVPE
14  bG1         IIAPCVLISSVAILIYFLPAKAGGQKCTVAINVLLAQTVFLFLVAKKVPE
15  mG1         IIAPCVLISSVAILIYFLPAKAGGQKCTVATNVLLAQTVFLFLVAKKVPE
16  cG1         IIVPCVLISSMAVLVYFLPAKAGGQKCTVSINVLLAQTVFLFLIAQKVPE
17  xG1         IIVPCVLISFVSILVYFLPAKAGGQKCTVSINILLAQTVFLFLVAQKIPE
18  tG1         IIAPCVLISSLVVLVYFLPAQAGGQKCTLSISVLLAQTIFLFLIAQKVPE
                II PCVLIS    L  FLPA AGGQKCT  I  LLAQT FLFL A K PE

19  bE1         IIVPCVLISGLVLLAYFLPAQAGGQKCTVSINVLLAQTVFLFLIAQKTPE
20  rE1         IIVPCVLISGLVLLAYFLPAQAGGQKCTVSINVLLAQTVFLFLIAQKIPE
                IIVPCVLISGLVLLAYFLPAQAGGQKCTVSINVLLAQTVFLFLIAQKIPE

21  bD1         ILVPCVLISFMINLVFYLPADCG EKTSMAISVLLAQSVFLLLISKRLPA
22  mD1         ILVPCVLISFMINLVFYLPGDCG EKTSVAISVLLAQSVFLLLISKRLPA
23  cD1         IVTPCVLIAFMAILVFYLPADSG EKMTLVISVLLAQSVFLLLVSQRLPA
24  tD1         FITPCVLISFLASLAFYLPAESG EKMSTAISVLLAQAVFLLLTSQRLPE
25  xD1         ILAPCVLIALMANLVFYLPADSG EKMTLAISVLLAQSVFLLLISQRLPE
26  tD1         FITPCVLISFLASLAFYLPAESG EKMSTAISVLLAQAVFLLLTSQRLPE
                   PCVLI     L FYLP   G EK    ISVLLAQ VFLLL   RLP

27  rA2         LIIPCLLISCLTVLVFYLPSECG EKITLCISVLLSLTVFLLLITEIIPS
28  cA2         LIIPCLLISCLTVLVFYLPSDCG EKITLCISVLLSLTVFLLLITEIIPS
29  rA3         LIIPCLLISFLTVLVFYLPSDCG EKVTLCISVLLSLTVFLLVITETIPS
30  cA3         MIIPCLLISFLTVLVFYLPSDCG EKVTLCISVLLSLTVFLLVITETIPS
31  rA4         LIIPCLLISCLTVLVFYLPSECG EKVTLCISVLLSLTVFLLLITEIIPS
32  cA4         LIIPCLLISCLTVLVFYLPSECG EKITLCISVLLSLTVFLLLITEIIPS
33  dAL         LIIPCVGISFLSVLVFYLPSDSG EKISLCISILLSLTVFFLLLAEIIPP
                 IIPC  IS L VLVFYLPS  G EK  LCIS LLSLTVFFL   EIIP

34  rB2         LIIPCVLITSLAILVFYLPSDCG EKMTLCISVLLALTVFLLLISKIVPP
35  cB2         LIIPCILITSLAILVFYLPSDCG EKMTLCISVLLALTVFLLLISKIVPP
36  rB3         LIIPCLGLSFLTVLVFYLPSDEG EKLSLSTSVLVSLTVFLLVIEEIIPS
37  rB4         LIIPCVLITSLAILVFYLPSDCG EKMTLCISVLLALTFFLLLISKIVPP
                LIIPC     L  LVFYLPSD G EK  L ISVL  LT FLL I  I P

38  gNA         LIIPCLGLSFLTVLVFYLPSDEG EKVSLSTSVLVSLTVFLLVIEEIIPS
39  dNA         LILPTVLISFLCVLVFYLPAEAG EKVTLGISILLSLVVFLLLVSKILPP
                LI P V  SFL VLVFYLP   G EKV L IS L SL VFLL    I P

40  hA1         TYLPCIMTVILSQVSFWLNRESVPARTVFGVTTVLTMTTLSISARNSLPK
41  bA1         TYLPCIMTVILSQVSFWLNRESVPARTVFGVTTVLTMTTLSISARNSLPK
42  bA2         TYLPCIMTVILSQVSFWLNRESVPARTVFGVTTVLTMTTLSISARNSLPK
43  bA3         TYLPCIMTVILSQVSFWLNRESVPARTVFGVTTVLTMTTLSISARNSLPK
                TYLPCIMTVILSQVSFWLNRESVPARTVFGVTTVLTMTTLSISARNSLPK

44  bB1         TYMPSTLITILSWVSFWINYDASAARVALGITTVLTMTTISTHLRETLPK
45  rB1         TYMPSTLITILSWVSFWINYDASAARVALGITTVLTMTTISTHLRETLPK
46  rB2         TYMPSILITILSWVSFWINYDASAARVALGITTVLTMTTINTHLRETLPK
47  rB3         TYMPSIMITILSWVSFWINYDASAARVALGITTVLTMTTINTHLRETLPK
                TYMPSIMITILSWVSFWINYDASAARVALGITTVLTMTTI THLRETLPK

48  rA1         MYIPSLLIVILSWISFWINMDAAPARVGLGITTVLTMTTQSSGSRASLPK

invariant          P                                            P
```

```
          351    <------------M3----------->          .          .
 1  hA1     TSSAVPLI GKYMLFTMVFVIASIIITVIVINTHH RSPST HVMPNWVR
 2  bA1     TSSAVPLI GKYMLFTMVFVIASIIITVIVINTHH RSPST HVMPEWVR
 3  mA1     TSSAVPLI GKYMLFTMVFVIASIIITVIVINTHH RSPST HIMPEWVR
 4  cA1     TSSAVPLI GKYMLFTMVFVIASIIITVIVINTHH RSPST HTMPPWVR
 5  s1A1
 6  s1A1
 7  xA1     TSSAVPLI GKYMLFTMVFVIASIIITVIVINTHH RSPST HTMPPWVR
 8  tA1     TSSAVPLI GKYMLFTMIFVISSIIITVVVINTHH RSPST HTMPQWVR
            TSSAVPLI GKYMLFTM FVI SIIITVVVINTHH RSPST H MP WVR

 9  hB1     TSLSVPII IKYLMFTMVLVTFSVILSVVVLNLHH RSPHT HQMPLWVR
10  bB1     TSLSVPII IKYLMFTMVLVTFSVILSVVVLNLHH RSPHT HQMPLWVR
11  mB1     TSLAVPII IKYLMFTMVLVTFSVILSVVVLNLHH RSPHT HQMPFWVR
12  tB1     TSLSVPII IRYLMFIMILVAFSVILSVVVLNLHH RSPNT HTMPNWIR
            TSL VPII I YLMFIM LV FSVILSVVVLNLHH RSP T H MP W R

13  hG1     TSQAVPLI SKYLTFLLVVTILIVVNAVVVLNVSL RSPHT HSMARGVR
14  bG1     TSQAVPLI SKYLTFLLVVTILIVVNAVVVLNVSL RSPHT HSMARGVR
15  mG1     TSQAVPLI SKYLTFLMVVTILIVVNSVVVLNVSL RSPHT HSMARGVR
16  cG1     TSQAVPLI GKYLTFLMVVTVVIVVNAVIVLNVSL RTPNT HSMSQRVR
17  xG1     TSTSVPLI VKYLTFLMVVTITIVANAVIVLNISL RTPNT HSMSSTVR
18  tG1     TSLNVPLI GKYLIFVMFVSMLIVMNCVIVLNVSL RTPNT HSLSEKIK
            TS  VPLI  KYL F   V   IV N VIVLN SL R P T HS

19  bE1     TSLSVPLL GRYLIFVMVVATLIVMNCVIVLNVSL RTPTT HAMSPRLR
20  rE1     TSLSVPLL GRYLIFVMVVATLIVMNCVIVLNVSL RTPTT HATSPRLR
            TSLSVPLL GRYLIFVMVVATLIVMNCVIVLNVSL RTPTT HA SPRLR

21  bD1     TSMAIPLI GKFLLFGMVLVTMVVVICVIVLNIHF RTPST HVLSEPVK
22  mD1     TSMAIPLV GKFLLFGMVLVTMVVVICVIVLNIHF RTPST HVLSEGVK
23  cD1     TSHAIPLI GKYLLFIMLLVTAVVVICVVVLNFHF RTPST HVMSDWVR
24  tD1     TALAVPLI GKYLMFIMSLVTGVIVNCGIVLNFHF RTPST HVLSTRVK
25  xD1     TSFAIPLI SKYLMFIMVLVTIVVVSCVIVLNLHF RTPST HAISERMK
26  tD1     TALAVPLI GKYLMFIMSLVTGVIVNCGIVLNFHF RTPST HVLSTRVK
            T    PLI G  L F M   T     C IVLN    RTP T H  S

27  rA2     TSLVIPLI GEYLLFTMIFVTLSIVITVFVLNVHH RSPST HNMPNWVR
28  cA2     TSLVIPLI GEYLLFTMIFVTLSIIITVFVLNVHH RSPST HTMPHWVR
29  rA3     TSLVIPLI GEYLLFTMIFVTLSIVITVFVLNVHY RTPTT HTMPTWVK
30  cA3     TSLVIPLI GEYLLFTMIFVTLSIVITVFVLNVHY RTPKT HTMPVWVR
31  rA4     TSLVIPLI GEYLLFTMIFVTLSIVITVFVLNVHH RSPRT HTMPAWVR
32  cA4     TSLVIPLI GEYLLFTMIFVTLSIIITVFVLNVHH RSPRT HTMPDWVR
33  dAL     TSLTVPLL GKYLLFTMMLVTLSVVVTIAVLNVNF RSPVT HRMAPWVQ
            TSL  PL  G YLLFTM LVTLS   T  VLNV   R P T H M  WV

34  rB2     TSLDVPLV GKYLMFTMVLVTFSIVTSVCVLNVHH RSPTT HTMAPWVK
35  cB2     TSLDVPLV GKYLMFTMVLVTFSIVTSVCVLNVHH RSPTT HTMPPWVR
36  rB3     SSKVIPLI GEYLLFIMIFVTLSIIVTVFVINVHH RSSSTYHPMAPWVK
37  rB4     TSLDIPLI GKYLLFTMVLVTFSIVTTVCVLNVHH RSPST HTMASWVK
             S   PL  G YL F M  VT SI   V V NVHH RS  T H M  WV

38  gNA     SSKVIPLI GEYLLFIMIFVTLSIIVTIFVINVHH RSSATYHPMSPWVR
39  dNA     TSLVLPLI AKYLLFTFIMNTVSILVTVIIINWNF RGPRT HRMPMYIR
             S V PLI   YLLF  I  T SI V    IN    R   T HRM    R

40  hA1     VAYATAMDWDWFIAVCYAFVFSALIEFATVNYF
41  bA1     VAYATAM  DWFIAVCYAFVFSALIEFATVNYFT          KRGYAWDG
42  bA2     VAYATAM  DWFIAVCYAFVFSALIEFATVNYF           TKRGWAW
43  bA3     VAYATAM  DWFMAVCYAFVFSALIEFATVNYF           TKRSWAW
            VAYATAM  DWF AVCYAFVFSALIEFATVNYF              W

44  bB1     IPYVKAI  DIYLMGCFVFVFLALLEYAFVNYIFFGKGPQKKGAGKQDQS
45  rB1     IPYVKAI  DIYLMGCFVFVFLALLEYAFVNYIFFGKGPQKKGASKQDQS
46  rB2     IPYVKAI  DMYLMGCFVFVFMALLEYALVNYIFFGRGPQRQKKAAEKAA
47  rB3     IPYVKAI  DMYLMGCFVFVFLALLEYAFVNYIFFGRGPQRQKKLAEKTA
            IPYVKAI  D YLMGCFVFVF ALLEYA VNYIFFG GPQ

48  rA1     VSYVKAI  DIWMAVCLLFVFSALLEYAAVNFV             SRQH

invariant
```

```
          401
 1  hA1    KVFIDTIPNIMFFSTMK
 2  bA1    KVFIDTIPNIMFFSTMK
 3  mA1    KVFIDTIPNIMFFSTMK
 4  cA1    KIFIDTIPNIMFFSTMK
 5  s1A1
 6  s1A1
 7  xA1    KIFIETIPNIMFFSTMK
 8  tA1    KIFIDTIPNVMFFSTMK
           K FI TIPN MFFSTMK

 9  hB1    QIFIHKLPLYLRLKRPK
10  bB1    QIFIHKLPLYLGLKRPK
11  mB1    QIFIHKLPPYLGLKRPK
12  tB1    QIFIETLPPFLWIQRPV
           QIFI  LP  L   RP

13  hG1    KVFLRLLPQLLRMHVRP
14  bG1    KVFLRLLPQLLRMHVRP
15  mG1    KLFLRLLPQLLRMHVRP
16  cG1    QVWLHLLPRYLGMHMPE
17  xG1    ELCLRTVPRLLRMHLRP
18  tG1    HLFLGFLPKYLGMQLEP
             L   P  L M

19  bE1    YVLLELLPQLLGSGAPP
20  rE1    QILLELLPRLLGLSPPP
             LLELLP LLG   PP

21  bD1    KLFLETLPEILHMSRPA
22  mD1    KFFLETLPKLLHMSRPA
23  cD1    GVFLEILPRLLHMSHPA
24  tD1    QIFLEKLPRILHMSRADE
25  xD1    EIFLNKLPRILHMSQPA
26  tD1    QIFLEKLPRILHMSRADE
             FL  LP  LHMS

27  rA2    VALLGRVPRWLMMNRPL
28  cA2    SFFLGFIPRWLFMKRPPLLLP
29  rA3    AVFLNLLPRVMFMTRPT
30  cA3    TIFLNLLPRIMFMTRPT
31  rA4    RVFLDIVPRLLFMKRPSVVKDNCRRLIESMHKMANAPRFWPEPVGEPGIL
32  cA4    RVFLDIVPRLLFMKRPSTVKDNCKKLIESMHKLTNSPRLWSETDMEPNFT
33  dAL    RLFIQILPKLLCIERPK
             F    P      RP

34  rB2    VVFLEKLPTLLFLQQPR
35  cB2    TLFLRKLPALLFMKQPQ
36  rB3    RLFLQRLPRWLCMKDP
37  rB4    ECFLHKLPTFLFMKRPG
             FL  LP  L    P

38  gNA    SLFLQRLPHLLCMRGN
39  dNA    SIFLHYLPAFLFMKRPRKTR
           S FL  LP    M

40  hA1
41  bA1    KSVVPEKPK
42  bA2    D
43  bA3    EG  KKVPEALEMKK


44  bB1    ANEKNKLEMNKVQVDAH
45  rB1    ANEKNKLEMNKVQVDAH
46  rB2    NANNEKMRLDVNKMDPH
47  rB3    KAKNDRSKSEINRVDAH
                          H

48  rA1    KELLRFRRK

invariant
```

```
          451          .          .          .          .          .
 1  hA1
 2  bA1
 3  mA1
 4  cA1
 5  s1A1
 6  s1A1
 7  xA1
 8  tA1

 9  hB1
10  bB1
11  mB1
12  tB1

13  hG1
14  bG1
15  mG1
16  cG1
17  xG1
18  tG1

19  bE1
20  rE1

21  bD1
22  mD1
23  cD1
24  tD1
25  xD1
26  tD1

27  rA2
28  cA2
29  rA3
30  cA3
31  rA4     SDICNQGLSPAPTFCNPTDTAVETQPTCRSPPLEVPDLKTSEVEKASPCP
32  cA4     TSSSPSPQSNEPSPTSSFCAHLEEPAKPMCKSPSGQYSMLHPEPPQVTCS
33  dAL

34  rB2
35  cB2
36  rB3
37  rB4

38  gNA
39  dNA

40  hA1
41  bA1
42  bA2
43  bA3

44  bB1
45  rB1
46  rB2
47  rB3

48  rA1

invariant
```

```
          501        .         .         .         .
 1  hA1                    RPSR                    EKQDKKIF
 2  bA1                    RPSR                    EKQDKKIF
 3  mA1                    RPSR                    DKQEKRIF
 4  cA1                    RPSR                    DKPDKKIF
 5  s1A1
 6  s1A1
 7  xA1                    RPSQ                    EKQPQKTF
 8  tA1                    RASK                    EKQENKIF
                           R S                      K   KIF

 9  hB1              PERDLMPEPPHCSSPGSGWG           RGTDEYFI
10  bB1             PERDQMQEPPSIAPRDSPGSGWG         RGTDEYFI
11  mB1              PERDQLPEPHHSLSPRSGWG           RGTDEYFI
12  tB1                TTPSPDSKPTIIS                RANDEYFI
                                                    R  DEYFI

13  hG1              LAPAAVQDTQSRLQNGSSGWSI         TTGEEVAL
14  bG1              LAPVAVQDAHPRLQNGSSSGWPI        TAGEEVAL
15  mG1              LAPAAVQDARFRLQNGSSSGWPI        MAREEGDL
16  cG1                EAPGPPQATRRRSSLGLM           VKADEYML
17  xG1              TDAAPPLAPLMRRSSSLGLM           MKADEYML
18  tG1               SEETPEKPQPRRRSSFGIM           IKAEEYIL
                                  R                    E  L

19  bE1                EIPRAASPPRRASSLGLL           LRAEELIL
20  rE1                EDPGAASPARRASSVGI            LLRAEELI
                       EDP AASP RRASS G             L   E

21  bD1                EDGPSPGTLIRRSSSLGYI          SKAEEYFS
22  mD1                EEDPGPRALIRRSSSLGYI          CKAEEYFS
23  cD1                ESPAGAPCIRRCSSAGYI           AKAEEYYS
24  tD1              SEQPDWQNDLKLRRSSSVGYI          SKAQEYFN
25  xD1              EPEPEPWSGVLLRRSSSVGYI          VKAEEYYS
26  tD1              SEQPDWQNDLKLRRSSSVGYI          SKAQEYFN
                                 RR SS GYI           KA EY

27  rA2             PPMELHGSPDLKLSPSYHWLETNMDAGEREETEEEEEEED
28  cA2            AEGTTGQYDPPGTRLSTSRCWLETDVDDKWEEEEEEEEEEE
29  rA3               SGEGDTPKTRTFYGAELSNLNCFSRADSKSCKEGYPC
30  cA3               SDEENNQKPKPFYTSEFSNLNCFNSSEIRCCKDGFVCQ
31  rA4    SPGSCPPPKSSSGAPMLIKARSLSVQHVPSSQEAAEDGIRCRSRSIQYCV
32  cA4    SPKPSCHPLSDTQTTSISKGRSLSVQQMYSPNKTEEGSIRCRSRSIQYCY
33  dAL             KEEPEEDQPPEVLTDVYHLPPDVDKFVNYDSKRFSGDYGI

34  rB2               HRCARQRLRLRRRQRE              REGEAVFF
35  cB2               QNCARQRLRQRRQTQERA            AAATLFLR
36  rB3                                               MDRFSF
37  rB4               LEVSLVRVPHPSQLHLATA            DTAATSAL

38  gNA                                              TDRYHYPE
39  dNA            LRWMMEMPGMSMPAHPHPSYGSPAELPKHISAIGGKQSKME

40  hA1
41  bA1                                               KVKDPLI
42  bA2                                              GKSVVNDK
43  bA3                                               KTPAVPT
                                                      K

44  bB1                GNILLSTLEIRNET                SGSEVLTG
45  rB1                GNILLSTLEIRNET                SGSEVLTG
46  rB2                ENILLSTLEIKNEM                ATSEAVMG
47  rB3                GNILLAPMDVHNEM                NEVAGSVG
                        NILL      NE                      G

48  rA1                                               RRHHKDD

invariant
```

```
          551        .         .         .         .         .
 1  hA1    TEDIDISDISG                                  KPGPPPMGFH
 2  bA1    TEDIDISDISG                                  KPGPPPMGFH
 3  mA1    TEDIDISDISG                                  KPGPPPMGFH
 4  cA1    AEDIDISEISG                                  KQGPVPVNFY
 5  s1A1
 6  s1A1
 7  xA1    AEEMDISHISG                                  KLGPAAVTYQ
 8  tA1    ADDIDISDISG                                  KQVTGEVIFQ
               DIS ISG                                  K

 9  hB1    RKPPSDFLFPK                                  PNRFQPELSA
10  bB1    RKPPNDFLFPK                                  PNRFQPELSA
11  mB1    RKPPSDFLFPK                                  LNRFQPESSA
12  tB1    RKPAGDFVCPVDN                               ARVAVQPERLF
           RKP  DF  P                                      QPE

13  hG1    CLPRSELLFQQWQRQG                            LVAAALEKLEK
14  bG1    CLPRSELLFRQRQRNGL                           VRAALEKLEKG
15  mG1    CLPRSELLFRQRQRNGL                           VQAVLEKLENG
16  cG1    WKARTELLFEKQKERDGL                          MKTVLEKIGRG
17  xG1    RKPRSQLMFEKQKERDGL                          MKVVLDKIGRG
18  tG1    KKPRSELMFEEQKDRH                            GLKRVNKMTSD
              R  LMF

19  bE1    KKPRSELVFE                                   QQRHRHGTWT
20  rE1    LKKPRRLVFEG                                  QRHRHGTWTA
            K    LVFE                                   Q

21  bD1    LKSRSDLMFEKQ                                SERHGLARRLT
22  mD1    LKSRSDLMFEKQ                                SERHGLARRLT
23  cD1    VKSRSELMFEK                                  QSERHGLASR
24  tD1    IKSRSELMFEKQSE                              RHGLVPRVTPR
25  xD1    VKSRSELMFEKQ                                SERHGLTSRAT
26  tD1    IKSRSELMFEKQSE                              RHGLVPRVTPR
            KSRS LMFEK

27  rA2    ENICVCAGLPDSSMG                               VLYGHGGLHLR
28  cA2    EEEEEEKAYPSRVPSGGSQ                        GTQCHYSCERQAGK
29  rA3    QDGTCGYCHHRRV                                 KISNFSANLTR
30  cA3    DMACSCCQYQRM                                  KFSDFSGNLTR
31  rA4    SQDGAASLADSKPTSSPTSLKARPSQLPVSDQASPCKCTCKEPSPVSPVT
32  cA4    LQEDSSQTNGHSSASPASQRCHLNEEQPQHKPHQCKCKCRKGEAAGTPTQ
33  dAL    PALPASHRFDLAAAGGISAHCFAEPPLPSSLPLPGADDDLFSPSGLNGDI

34  rB2    REGPAADPCTC                                  FVNPASVQGLA
35  cB2    AGARACACYAN                                   PGAAKAEGLN
36  rB3    PDGKESDTAV                                    RGKVSGKRKQ
37  rB4    GPTSPSNLYGSS                                  MYFVNPVPAA

38  gNA    LEPHSPDLKP                                    RNKKGPPGPE
39  dNA    VMELSDLHHPNCKIN                              RKVNSGGELGL
                                                            G

40  hA1
41  bA1    KKNNT                                         YAPTATSYT
42  bA2    KKEKASVMIQNN                                 AYAVAVANYA
43  bA3    KKTSTTFNIVG                                  TTYPINLAKD
           KK

44  bB1    VGDPKTTMYS                                    YDSASIQYR
45  rB1    VSDPKATMYS                                    YDSASIQYR
46  rB2    LGDPRSTMLA                                    YDASSIQYR
47  rB3    DTRNSAISFDN                                      SGIQYR
                                                             IQYR

48  rA1    EGGEGRFNFSA                                   YGMGPACLQ

invariant
```

```
          601
 1  hA1    SPLIKHP                    EVKSAIEGIKYIAETMKSDQESNN
 2  bA1    SPLIKHP                    EVKSAIEGIKYIAETMKSDQESNN
 3  mA1    SPLIKHP                    EVKSAIEGVKYIAETMKSDQESNN
 4  cA1    SPLTKNP                    DVKNAIEGIKYIAETMKSDQESSN
 5  s1A1
 6  s1A1
 7  xA1    SPALKNP                    DVKSAIEGIKYIAETMKSDQESNK
 8  tA1    TPLIKNP                    DVKSAIEGVKYIAEHMKSDEESSN
            P  K P                     VK AIEG KYIAE MKSD ES N

 9  hB1    PDLRRFIDGPNRAVALLP         ELREVVSSISYIARQLQEQEDHDA
10  bB1    PDLRRFIDGPNRAVGLPP         ELREVVSSISYIARQLQEQEDHDV
11  mB1    PDLRRFIDGPTRAVGLPQ         ELREVISSISYMARQLQEQEDHDA
12  tB1    SEMKWHLNGLTQPVTLPQ         DLKEAVEAIKYIAEQLESASEFDD
                   G    V L             L E    I Y A QL      D

13  hG1    GPELGLSQFCGSLKQAAP         AIQACVEACNLIACARHQQSHFDN
14  bG1    PESGQSPEWCGSLKQAAP         AIQACVEACNLIARARHQQTHFDS
15  mG1    PEVRQSQEFCGSLKQASP         AIQACVDACNLMARARRQQSHFDS
16  cG1    LESNRAQDFCQSLEEASP         EIRACVEACNHIANATREQNDFSS
17  xG1    MENNTSDDLVHSLNHAAP         EIRTCVEACCHIASATREKNDFKS
18  tG1    IDIGTTVDLYKDLANFAP         EIKSCVEACNFIAKSTKEQNDSGS
                             P          I  CV AC   A

19  bE1    ATLCQNLGAAAP               EIRCCVDAVNFVASSTRDQEATGE
20  rE1    AALCQNLGAAAP               EVRCCVDAVNFVAESTRDQEATGE
           A LCQNLGAAAP               E RCCVDAVNFVA STRDQEATGE

21  bD1    TARRPPAGSEQAQQELFS         ELKPAVDGANFIVNHMKDQNNYNE
22  mD1    TARRPPASSEQVQQELFN         EMKPAVDGANFIVNHMRDQNSYNE
23  cD1    VTPARFAPAATSEEQLYD         HLKPTLDEANFIVKHMREKNSYNE
24  tD1    IGFGNNNENIAASDQLHD         EIKSGIDSTNYIVKQIKEKNAYDE
25  xD1    PARVNPLNANNSQDQLYG         EIKPAIDGANFIVKHIRDKNDYNE
26  tD1    IGFGNNNENIAASDQLHD         EIKSGIDSTNYIVKQIKEKNAYDE
                        L               K   D  N IV      N Y E

27  rA2    AMEPETKTPSQASEILLS       PQIQKALEGVHYIADRLRSEDADSS
28  cA2    ASGGPAPQVPLKGEEVGSDQGLTLSPSILRALEGVQYIADHLRAEDADFS
29  rA3    SSSSESVNAVLSLSALSP         EIKEAIQSVKYIAENMKAQNVAKE
30  cA3    SSSSESVDPLFSFSVLSP         EMRDAIESVKYIAENMKMQNEAKE
31  rA4    VLKAGGTKAPPQHLPLSP         ALTRAVEGVQYIADHLKAEDTDFS
32  cA4    GSKSHSNKGEHLVLMSP          ALKLAVEGVHYIADHLRAEDADFS
33  dAL    SPGCCPAAAAAAAAADLSPTFEKPYAREMEKTIEGSRFIAQHVKNKDKFES
                                                 IA

34  rB2    GAFRAEPTAAGPGRSVGP       CSCGLREAVDGVRFIADHMRSEDDDQS
35  cB2    GYRERQGQGPDPPAPCGC         GLEEAVEGVRFIADHMRSEDDDQS
36  rB3    TPASDGERVLVAF               LEKASESIRYISRHVKKEHFISQ
37  rB4    PKSAVSSHTAGLPRDARLRSSGRFREDLQEALEGVSFIAQHLESDDRDQS
                                           A      I  H

38  gNA    GEGQALINL                    LEQATNSVRYISRHIKKEHFIRE
39  dNA    GDGCRRESESSDSILLSP         EASKATEAVEFIAEHLRNEDLYIQ
39  dNA    G                             AT  V  I  H   E

40  hA1                                    TKRGYAWDGKSVVPEKPKKV
41  bA1    PNLARGDP                    GLATIAKSATIEPKEVKPETKPPE
42  bA2    PNLSKDP                     VLSTISKSATTPEPNKKPENKPAE
43  bA3    TEFSAISKGA                  APSTSSTPTIIASPKTTCVQDIPT


44  bB1    KPMSSREGY                   GRALDRHGAHSKGRIRRRASQLKV
45  rB1    KPLSSREGF                   GRGLDRHGVPGKGRIRRRASQLKV
46  rB2    KAGLPRHSFG                  RNALERHVAQKKSRLRRRASQLKI
47  rB3    KQSMPKEGHG                  RYMGDRSIPHKKTHLRRRSSQLKI
           K                              R     K   RRR SQLK

48  rA1    AKDGIS                       VKGANNNNTTNPAPAPSKSPEEM

invariant
```

```
          651                <----------M4--------->
 1  hA1    AAAEWKYVAMV  MDHILLGVFMLVCIIGTLAVFAGRLIELNQQG
 2  bA1    AAEEWKYVAMV  MDHILLAVFMLVCIIGTLAVFAGRLIELNQQG
 3  mA1    AAEEWKYVAMV  MDHILLGVFMLVCLIGTLAVFAGRLIELHQQG
 4  cA1    AADEWKFVAMV  LDHLLLVIFMLVCIIGTLAVFAGRLIELNQQG
 5  s1A1
 6  s1A1
 7  xA1    ASEEWKFVAMV  LDHILLAVFMTVCVIGTLAVFAGRIIEMNMQE
 8  tA1    AAEEWKYVAMV  IDHILLCVFMLICIIGTVSVFAGRLIELSQEG
           A  EWK VAMV   DH LL  FM  C IGT  VFAGRLIE

 9  hB1    LKEDWQFVAMV  VDALFLWTFIIFTSVGTLVIFLDATYHLPPPDPFP
10  bB1    LKEDWQFVAMV  VDRLFLWTFIIFTSVGTLVIFLDATYHLPPADPFP
11  mB1    LKEDWQFVAMV  VDRLFLWTFIVFTSVGTLVIFLDATYHLPPPEPFP
12  tB1    LKKDWQYVAMV  ADRLFLYVFFVICSIGTFSIFLDASHNVPPDNPFA
           LK DWQ VAMV   D LFL  F    S GT  IFLDA    PP  PF

13  hG1    GNEEWFLVGRV  LDRVCFLAMLSLFICGTAGIFLMAHYNRVPALPFPGD
14  bG1    GNKEWFLVGRV  LDRVCFLAMLSLFVCGTAGIFLMAHYNRVPALPFPGD
15  mG1    GNEEWLLVGRV  LDRVCFLAMLSLFICGTAGIFLMAHYNQVPDLPFPGD
16  cG1    ENEEWILVGRV  IDRVCFFIMASLFVCGTIGIFLMAHFNQAPALPFPGD
17  xG1    ENEEWILMGRV  IDRVCFLVMCFVFFLGTIGTFLAGHFNQAPAHPFPGD
18  tG1    ENENWVLIGKV  IDKACFWIALLLFSIGTLAIFLTGHFNQVPEFPFPGD
               W L G     D  CF      F  GT  IFL  H N  P  PFPGD

19  bE1    EVSDWVRMGKA  LDSICFWAALVLFLVGSSLIFLGAYFNRVPQLPYPPC
20  rE1    ELSDWVRMGKA  LDNVCFWAALVLFSVGSTLIFLGGYFNQVPDLPYPPC
     1     E SDWVRMGKA  LDN CFWAALVLFSVGS LIFLG YFN VP LPYPPC

21  bD1    EKDCWNRVART  VDRLCLFVVTPIMVVGTAWIFLQGAYNQPPPQPFPGD
22  mD1    EKDNWNQVART  VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPLQPFPGD
23  cD1    EKDNWNRVART  LDRLCLFLITPMLVVGTLWIFLMGIYNHPPPLPFSGD
24  tD1    EVGNWNLVGQT  IDRLSMFIITPVMVLGTIFIFVMGNFNHPPAKPFEGD
25  xD1    EKDNWYRIART  VDRLCLFLVTPVMIIGTLWIFLGGAYNLPPSLPFPGD
26  tD1    EVGNWNLVGQT  IDRLSMFIITPVMVLGTIFIFVMGNFNHPPAKPFEGD
           E   W         DRL  F   P    GT  IF  G  N PP  PF GD

27  rA2    VKEDWKYVAMV  VDRIFLWLFIIVCFLGTIGLFLPPFLAGMI
28  cA2    VKEDWKYVAMV  IDRIFLWMFIIVCLLGTVGLFLPPYLAGMI
29  rA3    IQDDWKYVAMV  IDRIFLWVFILVCILGTAGLFLQPLMARDDT
30  cA3    IQDDWKYVAMV  IDRIFLWVFILVCILGTAGLFLQPLMTGDDM
31  rA4    VKEDWKYVAMV  IDRIFLWMFIIVCLLGTVGLFLPPWLAAC
32  cA4    VKEDWKYVAMV  IDRIFLWMFIIVCLLGTVGLFLPPWLAGMI
33  dAL    VEEDWKYVAMV  LDRMFLWIFAIACVVGTALIILQAPSLHDQSQPIDIL
              DWKYVAMV   DR FLW F   C  GT    L

34  rB2    VREDWKYVAMV  IDRLFLWIFVFVCVFGTVGMFLQPLFQNYTATTFLHP
35  cB2    VSEDWKYVAMV  IDRLFLWIFVFVCVFGTVGMFLQPLFQNYATNSLLQL
36  rB3    VVQDWKFVAQV  LDRIFLWLFLIASVLGSILIFIPALKMWIHRFH
37  rB4    VIEDWKFVAMV  VDRLFLWVFVFVCILGTMGLFLPPLFQIHAPSKDS
           V  DWK VA V   DR FLW F     LG    F   L

38  gNA    VVQDWKFVAQV  LDRIFLWTFLTVSVLGTILIFTPALKMFLRTPPPPSP
39  dNA    TREDWKYVAMV  IDRLQLYIFFIVTTAGTVGILMDAPHIFEYVDQDRII
              DWK VA V   DR  L IF  V   GT  I   A   F

40  hA1    KDPLFPN
41  bA1    PKKTFNSVSKIDRLSRIAFPLLFGIFNLVYWATYLNREPQLKAPTPHQ
42  bA2    AKKTFNSVSKIDRMSRIVFPVLFGTFNLVYWATYLNREPVLGVSP
43  bA3    ETKTYNSVSKVDKISRIIFPVLFAIFNLVYWATYVNRESAIKGMIRKQ
                 SVSK DK SRI FP LF  FNLVYWATY NRE

44  bB1    KIPDLTDVNSIDKWSRMFFPITFSLFNVVYWLYYVH
45  rB1    KIPDLTDVNSIDKWSRMFFPITFSLFNVVYWLYYVH
46  rB2    TIPDLTDVNAIDRWSRIFFPVVFSFFNIVYWLYYVN
47  rB3    KIPDLTDVNAIDRWSRIVFPFTFSLFNLVYWLYYVN
           IPDLTDVN IDRWSR  FP  FS FN VYWLYYV

48  rA1    RKLFIQRAKKIDKISRIGFPMAFLIFNMFYWIIYKIVRREDVHNK

invariant
```

```
          701
 1  hA1
 2  bA1
 3  mA1
 4  cA1
 5  s1A1
 6  s1A1
 7  xA1
 8  tA1


 9  hB1
10  bB1
11  mB1
12  tB1


13  hG1    PRPYLPSPD
14  bG1    PRSYLPSSD
15  mG1    PRPYLPLPD
16  cG1    PKTYLPP
17  xG1    SKLYQPST
18  tG1    PRKYVP
                P

19  bE1    M
20  rE1    IQP


21  bD1    PFSYLEKDKRFI
22  mD1    PFSYSEQDKRFI
23  cD1    PFDYREENKRYI
24  tD1    PFDYSSDHPRCA
25  xD1    PFIYTKEHRRLI
26  tD1    PFDYSSDHPRCA
           PF Y     R

27  rA2
28  cA2
29  rA3
30  cA3
31  rA4
32  cA4
33  dAL    YSKIAKKKFELLKMGSENTL


34  rB2    DHSAPSSK
35  cB2    GQGTPTSK
36  rB3
37  rB4


38  gNA
39  dNA    EIYRGK


40  hA1
41  bA1
42  bA2
43  bA3


44  bB1
45  rB1
46  rB2
47  rB3


48  rA1

invariant
```

Structure and Function

The following are the common features at the level of the derived amino acid sequence that have been used to delineate discrete regions of LGIC subunits:

1. A signal peptide that is removed upon membrane translocation of the polypeptide chain
2. An *N*-terminal extracellular agonist binding domain containing one or more sites for *N*-linked oligosaccharide attachment;
3. Three predicted transmembrane segments (termed M1, M2, and M3)
4. An intracellular region that often contains consensus sequences for regulatory sites of phosphorylation
5. A fourth predicted transmembrane segment (M4)
6. A short C-terminal region.

The Recognition Site for Agonists and Competitive Antagonists

The agonist and competitive antagonist binding site is of interest as a site to which pharmacologically active compounds may be targeted for therapeutic and research use. We have recently speculated that conservation in the LGIC receptors may be partly reflected in the structural similarities of the ligands for the different members of the LGIC family types (*see* Fig. 2a and 2b) (Cockcroft et al., 1990). A clear similarity in the structure of agonists is that they each contain a positively charged center (termed the positive pole), which is essential for activation as exemplified by studies on the nACh receptor for which tetramethylammonium is described as being the "minimal" agonist. In addition, they each have a π-electron system that contains an sp^2 hybridized electronegative atom, which induces a local dipole in the π-system and which can act as a hydrogen bond acceptor. The distance between the nitrogen atom of the positive pole and the electronegative atom is 4.5–5.5 Å for acetylcholine and GABA ligands, whereas for glycine, the distance is around 3.5 Å. The broad similarity of agonists is strikingly demonstrated by comparison of the almost totally rigid analogs cytisine and THIP, agonists of the nACh receptor and the $GABA_A$ receptor, respectively (*see* Fig. 2a). The similarity is more remarkable when it is considered that these receptors, which have opposite ion selectivity, are likely to be only distantly related in evolutionary terms.

What other neurotransmitters act on presently unidentified members of the LGIC superfamily? Clearly, glutamate (NMDA) is a good candidate (*see* Cull-Candy and Usowicz, 1987), since it has all the features of the proposed unified pharmacophore, as does histamine (*see* Hardie, 1989). Serotonin is accommodated less easily, since even though it has the π-system with an sp^2 electronegative atom this center cannot act as a hydrogen bond acceptor (but *see* Derkach et al., 1989 and Yakel and Meyer, 1988). The catecholamine transmitters do not fit the pharmacophore, since their aromatic rings do not contain electronegative atoms. Moreover, a local dipole is formed in the catechol moiety, which is opposite in orientation to that observed in the known LGIC agonists. This is, therefore, an indication that the catecholamine transmitters may not have undergone crossover from G-protein-coupled receptors during evolution to act on an LGIC receptor.

A useful conceptual framework for thinking about the agonist binding site is to consider it as two parts: (1) an essential region for receiving the message part of agonists, and (2) a region primarily involved in recognition of the address of agonists. The proposed similarity in agonist structures suggests that both these sites should be structurally well conserved, but with amino acid substitutions occurring at the region encoding the address to produce the pharmacological specificity of a given LGIC type. There is, however, no direct experimental evidence that the agonist binding sites in LGICs are so well conserved. Two other possibilities exist: (1) The overall position in the protein structure of the agonist binding site may be conserved, but the binding region itself (in terms of its main-chain conformation) could lack any structural correlation between LGIC types. This would be a situation analogous to

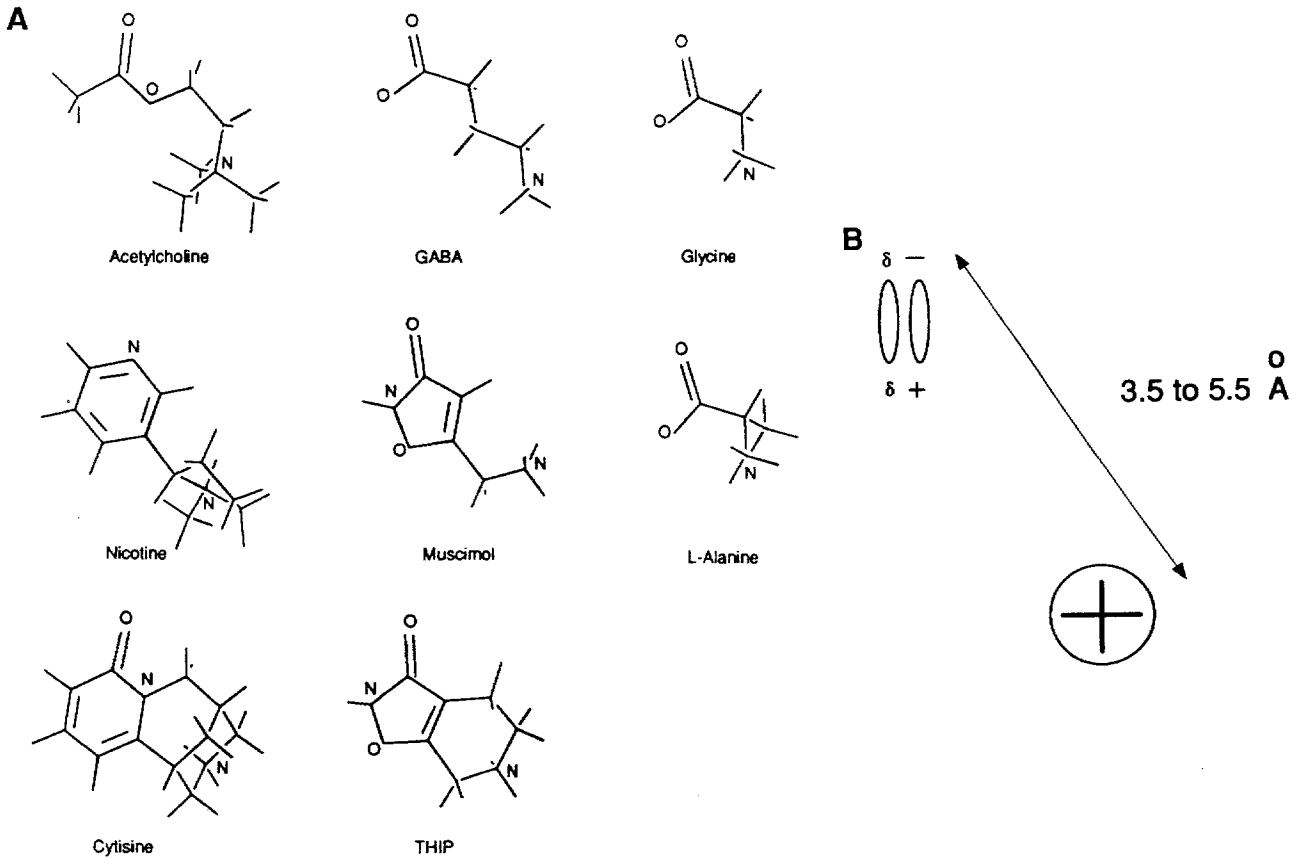


Fig. 2A. Structures of agonists of LGIC receptors. The ligands in columns from left to right are for the nACh receptor, the $GABA_A$ receptor, and the glycine receptor. In the rows from top to bottom are the neurotransmitters, semi-rigid agonist analogs and almost totally rigid agonist analogs.

Fig. 2B. Unified pharmacophore model for LGICs. The circle containing a plus symbol represents the positive pole. The local dipole is indicated by the symbols δ+ and δ-, with the latter representing the electronegative center of the local dipole.

that seen for the antigen combining sites of antibodies. (2) Even the overall location of the binding site may vary in the protein such that binding sites could have been made anew for the different LGIC types. Evolutionary arguments can be used to support either of these possibilities, but there is good precedence from the G-protein coupled receptors that receptor binding sites are well preserved in their overall structure.

The ability to be able to design highly potent rigid agonists for each of the different LGIC receptor types (Kanne and Abood, 1988; Krogsgaard-Larsen et al., 1983) may itself be an indication that their agonist binding sites are essentially similar. The implication of this is that activation can be achieved without any major change in the conformation of bound agonist and that an induced fit model for ligand recognition may be applicable. Another general feature of ligands supporting the notion of a structurally conserved binding site is that agonists tend to be small, whereas nonpeptide competitive antagonists are almost always large and tend to have molecular weights >200. (*see* Fig. 3). A simple model for agonist binding would then be that LGIC receptors undergo a change upon binding agonist that may be likened to the closing of a structural cleft round a substrate as seen for enzymes; antagonists may

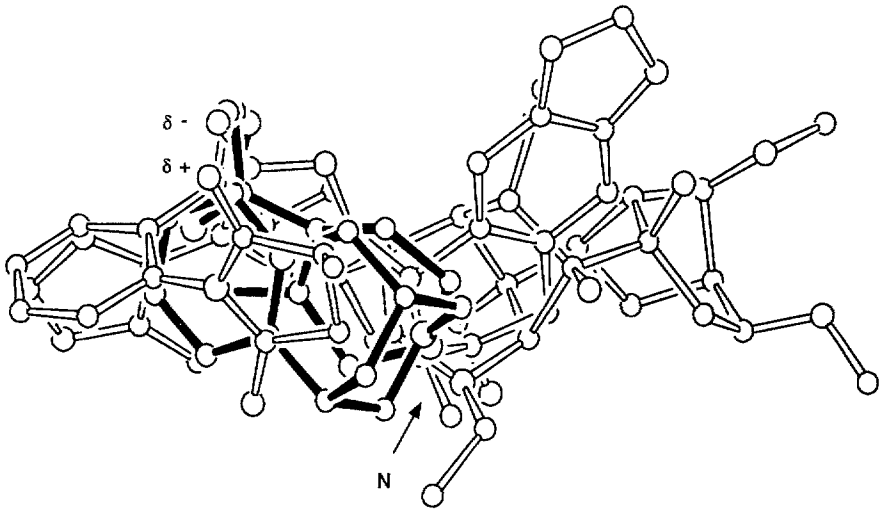


Fig. 3. Structural overlay of the rigid agonists cytisine and THIP (solid bonds), and the antagonists methyllycaconitine and bicuculline. The three centers of the molecules used for the superpositioning are the nitrogen atom of the positive pole, and the electronegative and electropositive centers of the local dipole.

bind and hold the cleft in the open conformation. In some cases, it has been proposed that a region of antagonists may mimic a portion of the agonist, thus explaining the competitive recognition that occurs. For acetylcholine, clear examples are dihydro-β-erythroidine and methyllycaconitine (MLA) (Wonnacott, 1987). The latter is interesting because the parent structure lycoctonine is related to aconitine which is known to act on voltage-gated sodium channels, but has no reported activity at nACh receptors. However, with the ester linkage present in MLA, a framework that can be fitted to acetylcholine is introduced (Ward et al., 1990). For GABA, the competitive antagonists bicuculline (Aprison and Lipkowitz, 1989) and pitrazepin (Boulanger et al., 1989) have been reported to contain a region with electronic properties resembling the carboxylate group of the endogenous transmitter. Strychnine and D-tubocurarine, competitive antagonists of the glycine receptor, and the nACh receptor, respectively, appear to be anomalous in that, although they contain a positively charged amine group, no other structural resemblance to their agonists is easily discernible.

We have recently proposed a general model for the agonist site of LGICs based on a highly conserved motif found in all subunits of all members of the LGIC superfamily (Cockcroft et al., 1990). This sequence, termed the *cys*-loop, is a 15-residue stretch that occurs at positions 128–142 in the α-subunit of the *Torpedo* nACh receptor (Luyten, 1986). The structure predicted for this region is a rigid, amphiphilic β-hairpin; at position 11 on its hydrophilic face there is an aspartic acid residue that is close to the turn of the loop (*see* Fig. 4). This aspartic acid residue is one of only two invariant acidic residue positions in the *N*-terminal extracellular region of all LGIC sequences and must, therefore, be a strong candidate for forming the anionic site. The residue at position 6 of the *cys*-loop is proposed as partly determining the specificity of agonist recognition. For the acetylcholine site, the residue at this position is a threonine, which is suggested to form a hydrogen bond with the ether oxygen of acetylcholine. For the glycine site, lysine occurs at this position, which could readily form an ion-pairing interaction with the carboxylate group of glycine. In the β-subunit of the $GABA_A$ receptor,

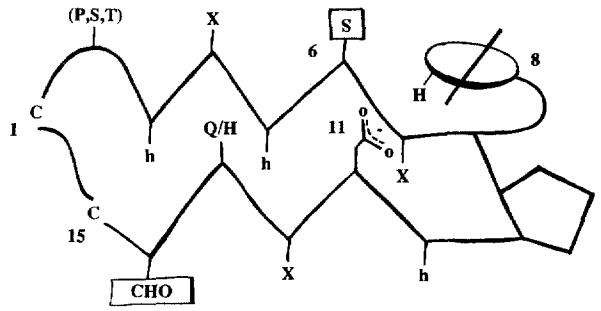


Fig. 4. A schematic representation of the *cys*-loop β-hairpin model. Only key residue positions are numbered. Abbreviations: boxed S = specificity residue at position 6; h = conserved hydrophobic residue positions; X = variable residue position; C, P, S, T, Q, H = one-letter amino acid rcde for amino acid residues; boxed CHO = site of *N*-glycosylation.

position 6 is an arginine, and this is proposed to form a potent hydrogen bond with a tyrosine at the turn of the *cys*-loop. This maintains the guanidinium group of the arginine at the right distance from the invariant aspartate so as to accommodate the carboxylate group of GABA. Experimental data on specific chemical modifications of the $GABA_A$ receptor and the glycine receptor are in accord with the proposal that the residue at position 6 of the *cys*-loop and other residues that are spatial neighbors of the invariant aspartate may make an important contribution to ligand specificity (*see* Table 1).

A feature of the *cys*-loop that makes it an attractive candidate for a primary determinant of the agonist binding site is that it is a rigid structure. This would reduce the entropy increase in binding owing to the loss of rotational degrees of freedom, compared to a flexible loop.

There have been many data to suggest that the region around residues 192–193 of α-subunits of nACh receptors is close to the agonist binding site. Labeling of these two cysteines after selective reduction of the disulfide bridge between them has been shown for agonist analogs such as bromoacetylcholine, and antagonists, such as MBTA and DDF (Langenbuch-Cachat et al., 1988; Dennis et al., 1988). It is interesting to note that whereas agonists protect labeling of the cysteines, D-tubocurarine and other antagonists do not (Karlin, 1980). DDF also affinity labels the Tyr-190, Trp-149, and Tyr-93 (Galzi et al., 1990) of the α-subunit of the *Torpedo* nACh receptor, indicating that the binding site is discontinuous in nature.

The snake toxins have proved to be powerful probes for the characterization of nACh receptors. The region around cysteines 192–193 of the α-subunit of the *Torpedo* nACh receptor has been shown to be a main determinant of α-bungarotoxin binding. By using synthetic peptides corresponding to regions of the α-subunit, high-affinity binding determinants have been located at the region 176–196 (Wilson et al., 1988) Recently, the α-sequence of the muscle nACh receptor from two snakes insensitive to α-bungarotoxin (Neumann et al., 1989) was shown to have undergone nonconservative substitutions around the 192–193 paired cysteines. A major change occurred at position 189, at which an asparagine residue was found to be a potential site of *N*-linked glycosylation. Other stretches of the amino acid sequence of the *Torpedo* α-subunit found to interact with α-bungarotoxin are: 1–16, 23–49, 100–115, 122–150 (Atassi et al., 1987), although using a solid-phase assay, the region α125–127 has been shown not to bind (Griesmann et al., 1990). α-Dendrotoxin has also been used to study the muscle-type nACh receptor. Evidence from simultaneous *N*-terminal sequencing indicates there to be four sites per oligomer, rather than two found for α-bungarotoxin (Conti-Tronconi and Raftery, 1986). Thus, regions other than 192–193, which is unique to the α-subunit, and that are common to each subunit can be expected to be involved in the binding of snake-toxins.

α-Bungarotoxin has also been shown to interact with the stretch 180–190 of the α5-subunit, tentatively identifying this subunit as a component of the α-bungarotoxin binding protein of neuronal tissues (McLane et al., 1990a). In the case of neuronal-bungarotoxin (also named κ-bungarotoxin [Chiappinelli et al., 1990]), which displays

Table 1
Proposed *Cys*-Loop Residues
Affected by Chemical Modification
of the GABA and Glycine Receptors
and Observed Effects on Ligand Binding

Chemical Reagent	LGIC subunit	*Cys*-loop residue	Observed effect
2,3-butadione[1]	GABA-β	Arg-6	GABA binding reduced
Phenylglyoxal[1]	GABA-β	Tyr-8	GABA binding reduced
Para-diazobenzene-sulfonic acid			
Tetranitromethane[2]			
N-acetylimidazone			
Diethyl-pyrocarbonate[3]	GABA-γ2	His-6	
		His-13	
	GABA-α	His-4	Benzodiazepine binding reduced
		His-13	
Fluorescein isothiocyanate[4]	GLY-α	Lys-6	Glycine binding reduced

[1]Widdows et al., 1987
[2]Maksay and Ticku, 1984
[3]Lambolez et al., 1987
[4]Gomez et al., 1989

selectivity for the neuronal forms of nACh receptors compared to the muscle-type, the region 51–70 of the α3-subunit of rat neuronal nACh receptor, which includes the motif WxDxxL conserved in all LGICs, was found to interact with neuronal-bungarotoxin as did the region 1–18 (McLane et al., 1990b)

Derivatized toxins have been used further to define interactions with the native receptor protein. The α-toxin of *Naja naja siamensis* was fluorescence labeled at lysines at positions 23, 35, 49, and 69, allowing energy transfer experiments to study the orientation and interaction of this toxin with the *Torpedo* nACh receptor (Johnson et al., 1990) These labeled residues were found not to be part of the binding surface, the major axis of the neurotoxin was tilted in a perpendicular projection from the membrane, and the receptor binding site was estimated to be 40 Å from the lipid membrane surface. The receptor-toxin complex has also been studied by use of photoactivatable derivatives of toxin-α from *Naja nigricollis* with reactive moieties at Lys-15, Lys-47, and Lys-51 (Chatrenet et al., 1990). At the high-affinity toxin binding site, toxin-α Lys-15 labeled predominantly the α-subunit, whereas Lys-51 reacted with the δ. For the low-affinity site, toxin-α Lys-47 labeled the α- and β-subunits, whereas Lys-15 and Lys-51 labeled γ and δ. In accord with these results, a coexpression study in which α-bungarotoxin binding was measured showed that the α–δ combination gave rise to a high-affinity site, whereas for α–γ a low-affinity binding was obtained (Kurosaki et al., 1987).

Lophotoxin gives further insights into the binding cavity of nACh receptors, because it differs in structure so much from the classical ligands (Wonnacott, 1987). The covalent attachment of lophotoxin to the tyrosine at position 190 of the *Torpedo* nACh receptor (Abramson et al.,

1988,1989) by reaction with the epoxide groups of the ligand may be made facile by the acidic residue at the anionic site. Neosurugatoxin is another natural product that acts as an antagonist of nACh receptors but with higher potency at neuronal forms than for the muscle-type receptor (Rapier et al., 1990; Wada et al., 1989; Lukas, 1989; Luetje et al., 1990). Interestingly, the debromination of neosurugatoxin results in a 100-fold decrease in binding affinity, whereas removal of the sugar moieties is much less dramatic (approx threefold loss) (Yamada et al., 1987). It could be that the bromine atom forms favorable soft–soft interactions with the sulfur atoms of the 192–193 disulfide bridge.

Recent experiments indicate that, when antagonists are bound at the nACh receptor, part of the binding site is formed by the interface between subunits. It has been shown that *d*-tubocurarine photoaffinity labels the γ- and δ-subunits of the *Torpedo* nACh receptor, in addition to the α-subunit (Pedersen et al. 1990). The IC_{50} for inhibition of specific labeling of the γ-subunit (40 n*M*) and the δ-subunit (0.9 μ*M*) gave good correspondence to the binding constants of *d*-tubocurarine at high- (35 n*M*) and low-affinity sites (1.2 μ*M*) of the *Torpedo* nACh receptor. In accord with this, coexpression of the subunit combinations α–γ and α–δ in fibroblasts resulted in high- and low-affinity *d*-tubocurarine sites, respectively (Blount and Merlie, 1988). These studies indicated that the two types of binding site may be formed at the α–γ and α–δ interfaces. They were also taken to suggest that the clockwise arrangement of subunits in the muscle-type receptor is α–γ–α–δ–β. However, in an electron microscopy study using probes for the α, β and δ-subunits of the *Torpedo* nACh receptor, the arrangement of subunits was found to be α–β–α–γ–δ (Kubalek et al., 1987). Interestingly, this latter arrangement would explain the ability of the β2-subunit of neuronal nACh receptors to substitute for the β-subunit in the muscle receptor, since the β-subunit would be flanked by two highly conserved α-subunits.

Other Features of the Extracellular Domain

The main immunogenic region (MIR), to which >60% of the antibodies in myasthenic serum bind, is a conformation-dependent epitope of the extracellular region of muscle-type nACh receptors. A continuous component of the MIR has been mapped, using overlapping synthetic peptides, to the region 67–76 of the α-subunits of the human muscle and *Torpedo* nACh receptor (Tzartos et al., 1990). Recently, the point mutations α68N→D and α71D→K (Saedi et al., 1990) indicated these positions to be important, in accord with a study using peptides in which substitutions by glycine were made (Bellone et al., 1989). Other regions of the MIR reported to bind antibodies are the stretches 1–14, 25–36, 41–53, 102–114, 128–138,172–182, and 188–198 (Mulac-Jericevic et al., 1987).

Using antibodies raised to short synthetic peptides, it was shown that the sequence stretches α81–85, α127–132, and α190–195 were freely accessible and presumed to be at the surface of the receptor (Maelicke et al., 1989). For the α1-subunit of the $GABA_A$ receptor, a similar approach indicated the *N*-terminus and *C*-terminus are accessible in the native $GABA_A$ receptor (Duggan and Stephenson, 1989). In the gene for the α-subunit of the human muscle receptor, a novel exon leads to an insertion of 25 residues between positions 58 and 59. In no other LGIC subunit sequences is there such a sizable insertion generating an additional isoform and presumably this region forms an additional surface loop structure (Beeson et al., 1990).

It has been reported for the *Torpedo* nACh recep-tor that binding of agonist causes the displacement of 4–5 calcium ions per receptor oligomer (Chang and Newmann, 1976). Interestingly, a sequence match to the EF-hand calcium-binding motif (Godzik and Sander, 1989) is conserved in all nACh receptor subunits over the stretch corresponding to 93–104 of the *Torpedo* α-subunit (*see* Fig. 5). Typically, the flanking sec-

Calcium Binding Site

	1	2	3	4	5	6	7	8	9	10	11	12
	X		Y		Z		Y		X			Z
Parvalbumin	D	S	D	G	D	G	K	I	G	V	D	E
Calmodulin	N	I	D	G	D	G	E	V	N	Y	E	E
Glucose-BP	D	L	N	K	D	G	Q	I	Q	I	-	E
Troponin-C	D	K	N	N	D	G	R	I	D	F	D	E
nAChR (Mouse)												
alpha	Y	N	N	A	D	G	D	F	A	I	V	K
beta	L	N	N	N	D	G	N	F	D	V	A	L
gamma	E	N	N	V	D	G	V	F	E	V	A	L
delta	E	N	N	N	D	G	S	F	Q	V	S	Y
Seq. Motif	O	x	O	x	O	G	=O	h	O	x	i	O

Quicho et al. CSH Symposium 1987

Fig. 5. Calcium binding motif. Numbering refers to positions in the EF–hand motif (Godzik and Sander, 1989). Abbreviations: X, Y, and Z = coordination axis of oxygen atoms; (G = O) = the carbonyl oxygen atom of glycine required for calcium coordination; O = positions of the motif able to donate a side-chain oxygen atom for coordination; h = conserved hydrophobic residue position; x = variable residue position.

ondary structure of such calcium binding sites is α-helix, but this may not be an essential requirement. The number of side-chain oxygen atoms for chelation, which in the different subunits ranges from 3–6, may determine the binding affinity of the metal ion. The motif is absent in the anionselective members of the LGIC superfamily, where instead a histidine-rich segment occurs, except in an α-subunit of the human $GABA_A$ receptor, in which case a deletion of the polypeptide chain is evident.

The Transmembrane Domain and the Ion Channel

The known LGIC subunits have the common feature of four hydrophobic segments, each of which is of an appropriate length to span the membrane in an α-helical conformation with 6 or 7 helical turns. These transmembrane segments are termed M1 through M4 in order of their appearance in the polypeptide chain, and occur at equivalent positions in each of the known receptor subunits. M1, M2, and M3 are always closely linked, being separated by short, hydrophilic segments (i.e., <8 residues), and M1 starts at about 200 residues in from the *N*-terminus. M4 is close to the C-terminus and is separated from the M1–M3 cluster by a hypervariable region, termed the major intracellular domain.

Initially, in the case of the nACh receptor, an additional transmembrane segment, an amphipathic helix termed MA was proposed to form the ion channel wall with its hydrophilic charged face (Finer-Moore and Stroud, 1984). However, the construction of mutants of the α-subunit of the *Torpedo* nACh receptor in which MA was deleted indicates that this segment is not essential for forming the gated ion channel response (Mishina et al., 1985). Moreover, an MA equivalent is not present in the subunits of the other members of the LGIC superfamily. Models incorporating MA in the membrane have been largely abandoned, and MA (now termed HA, an amphiphilic helix) is now thought to be located cytoplasmically. It is of note that, as yet, there is no function assigned to HA, even though it is well conserved in muscle and neuronal nACh receptor subunits, and particularly so in α-subunits.

The transmembrane arrangement of M1–M4 places the C-terminus on the extracellular side of the membrane. Indeed, using a hydrophilic reducing reagent (Dunn et al., 1986), it was reported that the disulfide linkage between oligomers of the *Torpedo* nACh receptor was on the extracellular side. More recently, it has been shown that this link is between δ-subunits of adjacent oligomers (DiPaola et al., 1989). This, therefore, supports the model of membrane topology of LGIC subunits in which the *N*-terminal domain is extracellular, there are four transmembrane segments (M1–M4), and the C-terminus is also extracellular.

Analysis of the photoreaction center, a transmembrane protein for which a structure (Deisenhofer et al., 1985) and several related sequences are known, reveals a higher degree of conservation for the contacts between one transmembrane

helix with its neighbors than for sites on the helix facing the lipid bilayer (Henderson et al., 1988; Rees et al., 1989). The conservation of the four transmembrane segments of LGIC subunits is M1 > M2 > M3 > M4 (*see* Fig. 1). By analogy, therefore, this pattern of conservation suggests that M1 may be packed towards the middle of a bundle of transmembrane helices, making extensive intra- and intersubunit contacts, and that M4 is on the outside and exposed to the lipid (*see* Fig. 6). This is consistent with mutagenesis experiments, in which foreign transmembrane segments (from interleukin-2 receptor and vesicular stomatitis virus glycoprotein) were shown to replace M4 of the α-subunit of the *Torpedo* nACh receptor without loss of channel activity, whereas similar replacement of M1, M2, or M3 resulted in loss of activity (Tobimatsu et al., 1987).

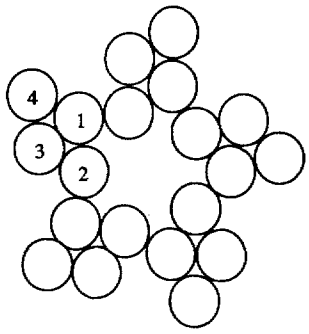


Fig. 6. Schematic model of the proposed bundling of transmembrane α-helices in a receptor oligomer. The numbers are labels for the transmembrane segments M1–M4 of a single subunit.

There is much evidence from experiments on muscle-type nACh receptor to suggest that M2 is an important determinant of the ion channel. Several such studies have made use of different noncompetitive antagonists that block the open channel. The neuroleptic chlorpromazine, which can be used as a photoaffinity reagent, was shown to label the serine residue at positions 262 and 254 of the δ-subunit and the β-subunit and a leucine residue at position 257 of the β-subunit of the *Torpedo* nACh receptor (Giraudat et al., 1987). The serines of the β- and δ-subunits are homologous sites that are positioned a third of the way into the M2 sequence from its cytoplasmic end (position 330 in the alignment in Fig. 1). Triphenylmethylphosphonium (TPMP) also labels this site and, additionally, the equivalent site in the α- and β-subunits (Hucho et al., 1986). Recently, the importance of a serine residue at this position has been demonstrated using site-directed mutagenesis and expression of altered receptors in the *Xenopus* oocyte system (Leonard et al., 1988; Charnet et al., 1990). Decreasing the number of the serine residues at the homologous sites in the mouse receptor led to a reduction in the equilibrium binding of QX-222, a derivative of lidocaine, and to marked changes in ion channel properties. These findings provide strong support for the suggestion that M2 forms part of the pore of the channel and that the serines contribute to the binding site of the channel-blocking noncompetitive antagonists. Interestingly, synthetic peptides with a high serine content that resemble the M2 sequence and that have an α-helical conformation have been shown to form ion channels with permeability and lifetime characteristics that resemble the channels of nACh receptors (Lear et al., 1988).

In an earlier series of experiments Numa's group showed that the δ-subunit M2 region and flanking sequence were chiefly responsible for observable differences in ion conductance of the *Torpedo* electric organ and bovine muscle forms of nACh receptor (Imoto et al., 1986). Using site-directed mutagenesis of the subunits of the *Torpedo* receptor, three important sites (1, 2, and 3; *see* Fig. 1) that lie at the ends of M2 have been identified (Imoto et al., 1988). This is in contrast to other sites possessing charged residues and in the vicinity of M2, where changes introduced had no effect on ion conductance. Remarkably, an almost linear inverse relationship is seen between channel conductance and the net negative charge carried at the above three sites. Changes at site 2 have a stronger effect than changes at the other two sites. In addition, evidence is provided that

magnesium ions interact with negatively charged residues at position 1 (cytoplasmically located) and position 3 (extracellularly located) to reduce selectively outward and inward currents, respectively. In contrast, changes that decreased the net negative charge at position 2 displayed reduced sensitivity to magnesium for both inward and outward currents. These observations of Imoto et al. led to the proposal that each of these positions is at or close to the mouth of the ion channel, and contributes to rings ("Imoto rings") of negative charge that selectively repel anions and concentrate cations ready for passage through the channel, and that position 2 may be close to or at the constriction of the ion channel.

The studies carried out so far on the muscle-type nACh receptor suggest two aspects of the channel that are important for ion passage. The first is rings of negative charge at either end of M2 and the second is the presence of polar hydroxyl containing residues towards the middle of the pore. Therefore, it should be possible to see whether, in anion selective members of the superfamily, changes at these sites can explain the observed switch in ion selectivity.

In comparing the residues at "Imoto ring" positions and the surrounding sequence of cation and anion selective channels, it can be seen that there is no correlation between the overall ring charge and the type of ion to flow through the ion channel (but see "Imoto ring" position 4). However, electrostatic interactions are long-range forces that could act over distances greater than the diameter of a single helix, and therefore absolute positioning of appropriately charged residues may not be essential; thus, it may be necessary to search for analogous "ring" residues on M1 and/or M3.

A major difference between anions and cations apart from their charge is the way in which they coordinate water molecules. Whereas it is the oxygen atom of water molecules that is involved in coordinate bonding of cations, with anions, waters interact via hydrogen bonds. From experiment, it is proposed that at least some of the inner solvation waters are lost during passage of ions through both cation and anion selective channels, since the minimum bore diameters are about 7.5 and 5.5 Å for cation- and anion-selective channels, respectively (Bormann et al., 1987). If this is the case, then the importance of the serine and threonine residues might be in forming an appropriate interaction with the migrating ion in the cation-selective and anion-selective channels, respectively. In the case of cation selective channels, the serine hydroxyl groups might be hydrogen bonded to the main chain carbonyl groups at position (i–3) or (i–4) of the M2 helix as is commonly suggested for this residue in α-helical structures. This would present to the channel an oxygen atom for coordination to cations. Threonine residues present in the anion-selective channels, however, may be prevented from forming such hydrogen bonds by means of steric restrictions involving the side-chain methyl group, which would then leave their hydroxyl groups free to hydrogen bond with anions passing through the channel. This difference would also account for the size difference of the two types of channel. At present, intuitive electrostatic arguments suggest that, in general, the pore of an ion channel will not contain charged residues: these would interact so strongly with a migrating ion that either the channel would be totally blocked or migration would be much too slow.

We have commented previously on some of the similarities that are seen between agonists and the proposed agonist binding sites among members of the LGIC family. It has also been observed that several channel blocking agents are effective on more than one of the cation channel members or candidate members of the LGIC superfamily. Thus, phencyclidine and MK801, in addition to their well-known effects on NMDA receptors, also act on nACh receptors (Albuquerque et al., 1980; Ramoa et al., 1990; Galligan and North, 1990; Kavanaugh et al., 1989). Mecamylamine and neuroleptics normally considered as nicotinic channel agents are effective on NMDA receptors (O'Dell and Christensen, 1988; Reynolds and Miller, 1988). Mg^{2+} ions (Huettner and Bean, 1988; Neher and Steinbach, 1978) are potent blockers

of both. This "crossover" does not seem to extend to the anion-selective channels of the LGIC receptors. However, picrotoxin, a channel blocker of the $GABA_A$ receptor, has been found to act on the glycine receptor (Akagi and Miledi, 1988). This suggests that the region of LGICs defining the ion channel is distinctive for each of the two types (i.e., anion selective and cation selective), but is conserved within each of these types. In accord with this, although M2 is not as conserved as M1 between the nACh receptor and the $GABA_A$ receptor, within each type of receptor it is the most conserved of the transmembrane segments.

The two types of ion channel do, however, show some similarities when examined electrophysiologically. Both show weak ion selectivity, maximum conductances of about 80 pS, average conductances of about 30 pS, and multiple conductance states i.e., nACh receptor: 9, 20, 30, 40 pS, $GABA_A$ receptor: 12, 20, 30, 46 pS (Bormann et al., 1987).

There are two unrelated proteins that are said to display partial sequence similarity to members of the LGIC superfamily. Kosower (1988) has proposed that the region preceding M4 in the $GABA_A$ receptor resembles a segment in the anion-exchange protein and that this is because of a functional requirement for anion transfer across the membrane. The suggestion is that this region in the $GABA_A$ receptor is functionally equivalent to MA of the nACh receptor. However, this region is not conserved among $GABA_A$ receptor subunits. Therefore, this suggestion does not seem to hold in the light of evidence that M2, and not MA forms, the pore of the ion channel of LGICs. It has also been suggested that there is a resemblance between transmembrane segments of the ryanodine receptor and M1, M2 and M3 segments of the nACh receptor. However, this seems less likely in the light of the cloning of the inositol trisphosphate receptor, which shows distinct homology with the ryanodine receptor, but the initially proposed M2 and M3 segments are not conserved (Furuichi et al., 1989).

The Major Intracellular Domain

The major intracellular domain is highly variable both in length and in sequence. It ranges in size from approx 100–250 residues, with deletion mutagenesis experiments indicating that a length as short as 80 residues does not abolish function (Mishina et al., 1985). The differences in the position and number of introns over this region suggest that intron slippage and the conversion of intronic sequence into coding region are partly the cause of the length variation (Nef et al., 1988). Its low sequence conservation is evident by comparisons of cognate subunits in different species and suggests that it lacks a conserved folded protein domain. For the *Torpedo* nACh receptor, 28% random coil is observed by CD spectroscopy of the whole receptor (Mielke and Wallace, 1988), which is similar in amount to the proportion that this region would represent of the oligomeric protein.

The lack of conservation in this region is surprising, because its cytoplasmic location suggests that it might make important interactions with cytoskeletal components. Such interactions could be involved in the localization of the receptors at their cellular sites of function. A 43-kD protein (Carr et al., 1987 that is myristilated at its *N*-terminus (Carr et al., 1989) and has a conserved cAMP-dependent phosphorylation site (Frail et al., 1988) is tightly associated with the muscle-type of nACh receptor and has been shown to interact with the β-subunits of neighboring oligomers (Burden et al., 1982) in the formation of aggregates of the receptor (Lo et al., 1980; Cataud et al., 1981). However, fibroblasts and other nonexcitable cell types also contain the 43-kD protein (Musil, 1989). Whether extrinsic proteins serving a similar role are associated with other LGICs is not known, although a 93-kD extrinsic protein has been identified for the glycine receptor (Langosch et al., 1988).

The results of mutagenesis deletion experiments suggest that the main intracellular domain does not play a significant role in the ligand-gated functioning of the receptors (Mishina et al., 1985).

It does, however, contain potential sites for serine/threonine and tyrosine phosphorylation, which may be involved in the enhancement of desensitization of the receptors (Qu et al., 1990). It is PEST-rich (meaning that it has a high content of proline, glutamate, serine, and threonine), which may predispose it to enzymatic degradation (Bachmair et al., 1986). The region shown to be susceptible to proteolytic cleavage includes HA (Roth et al., 1987; Dwyer, 1988). Interestingly, the region just beyond M3 in LGICs is moderately conserved and shows some similarity to a motif identified for the transferrin receptor that in this case is involved in receptor internalization (Jing et al., 1990).

Quaternary Structure

The proposed pentameric form of the LGICs was initially established by stoichiometric analysis by simultaneous *N*-terminal sequencing of whole oligomers of the *Torpedo* nACh receptor (Raftery et al., 1980). However, estimates of subunit stoichiometries by *N*-terminal sequencing are not definitive proof that the *Torpedo* receptor is pentameric. This is because the extent of *N*-terminal block by acetylation of the free amino terminus may vary for different types of subunits and may depend on the type of amino acid at their *N*-termini. Serine is the most prevalent of the amino acids to give rise to amino-terminal acetylation (Persson et al., 1985). It is therefore noteworthy that the terminal residue of the α- and β-subunit of the *Torpedo* receptor is serine, whereas for the γ- and δ-subunits it is glutamate and valine, respectively. Thus, estimates of the levels of the α- and β-subunit may be underestimated.

Electron microscopy has shown the overall shape of the *Torpedo* nACh receptor, including high-density regions corresponding to each of the subunits that are interpreted in terms of a pentameric structure (Toyoshima and Unwin, 1988; Mitra et al., 1989). Protein chemical analysis of the glycine receptor is in accordance with a pentameric oligomer (Langosch et al., 1988), whereas for the $GABA_A$ receptor a tetrameric form (Stephenson, 1988) has been proposed, and for the brain nACh receptor, the possibility of it being a tetramer has not been excluded (Lindstrom et al., 1987).

Evolutionary Diversity of LGICs

It seems that present-day LGIC receptors have probably arisen as a result of divergent evolution from a common ancestor. It is clearly more likely for a preexisting protein to become adapted to perform a modified, but still basically similar function than it is for the protein to be evolved *de novo*. Sufficient information is now at hand to permit a reasonable proposal of how the superfamily may have evolved. This is of interest because it may lead to insights into the way in which LGIC receptors function and how they are integrated into the overall physiology of complex nervous systems.

Origins of the Superfamily

The evolutionary tree shown in Fig. 7 shows branch points representing both divergence of subunits from one another and divergence of species within individual subunit groupings.

As with all such estimates of evolutionary history, both the branching structure and the dating of the tree should be regarded as tentative. The positions of the branches leading to the *Drosophila* sequences are particularly uncertain: doubtless, this uncertainty will be reduced as more insect sequences (and those of other invertebrates) become available.

Under the assumption of the analysis and using the calibration of the time scale as described in the legend of Fig. 7, the initial branch point off the right-hand end of the diagram would have been at least 2000 million years ago. This would roughly correspond with, or exceed, current estimates for the time of origin of eukaryotes. This date for the common origin of the receptors is, therefore, surprisingly early. It strongly suggests

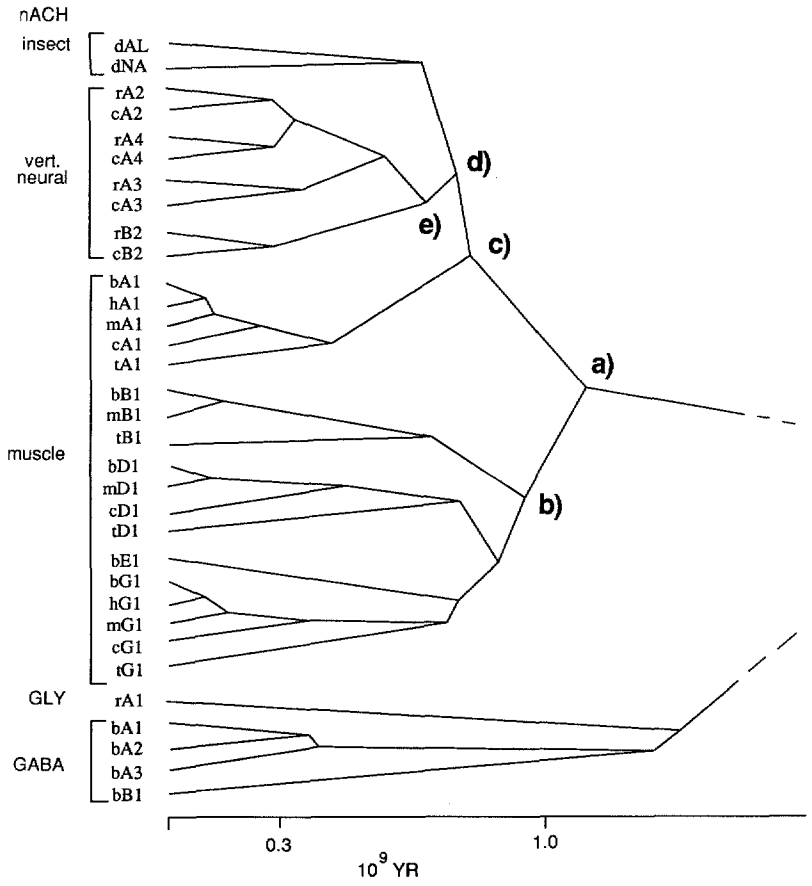


Fig. 7. Evolutionary, tree computed from the aligned sequences. The nomenclature scheme referring to the sequences is as described in the legend of Fig. 1. The sequence data were analyzed at both the nucleotide and protein levels, as complete sequences, and split into functional regions. In addition, at the nucleotide level, each of the codon positions was analyzed separately. The techniques of analysis have been described previously (Bishop and Friday, 1985). In constructing a time scale, the assumption has been made that the overall rate of change has been stochastically constant over the range of the tree. Calibration from relative to absolute units of time has been made using evidence from the fossil record for the time of divergence of the lineages leading to mammals and birds at approx 300 million years ago. This point is represented on the tree a number of times for the divergence of the lineage to chicken within the different subunits.

that the ancestral protoreceptor originated in a unicellular organism and raises the possibility that members of this structurally related protein set might be widely dispersed throughout living organisms. This may include plant and fungal tissues, eukaryotic viruses, and any of the non-nervous cell-types of animals. However, such proteins could serve biological roles other than those typically seen for the presently recognized group of LGICs.

It can only be speculated what the role of the protoreceptor was, although in a unicellular organism, it could have been osmotic regulation. The organism containing this early form of protein may possibly have been a prokaryote, since evidence from electrophysiological studies indicates that ion channels with weak ion selectivity occur in *Escherichia coli* and yeast (Saimi et al., 1988). Among eukaryotes, a glutamate receptor with a cation-selective ion channel is present in

Paramecium, and hydrozoans display responses with a nicotinic-type pharmacology.

Since all the receptor proteins derived from the root of the tree show ligand gating, it appears that the gating mechanism developed before the formation of separate lines of evolution for the cation-selective and anion-selective LGICs. The most likely candidates for the activating ligand of the earliest receptors are glutamate and glycine, since these are essential cellular metabolites. Perhaps an early organism used primitive forms of LGIC receptors in seeking out nutrient-rich environments. Nevertheless, the specialization of ion channel selectivity appears to have been a very early step occurring well before hetero-oligomerization events for any of the presently established LGICs.

Events in nACh Receptor Evolution

The large number of sequences of nACh receptor subunits from a variety of tissue types and species permits a much more detailed analysis of this protein than is yet the case for other members of the superfamily. The part of the tree concerned with the evolution of nACh receptors indicates evolution from a deduced ancestral homo-oligomer to a hetero-oligomeric form (Fig. 7, branch-point "a") not yet differentiated into muscle and neuronal types. The date for this duplication event is estimated to be 900 to 1200 million years ago. It remains uncertain, therefore, whether this duplication took place before or after the formation of metazoa. If the homo-oligomer was present in a single-celled eukaryote, this organism may have been a simple sensorimotor unit, responding to environmental stimuli in a manner like that of the myoepithelial cells of *Hydra*. It is also not known if the ancestral homo-oligomer was an acetylcholine receptor/cation channel. It seems more likely that some of the present-day glutamate receptors evolved from the primitive ion channel, and that differentiation to an ACh-recognizing type occurred subsequently.

The initial branch of the common lineage of the non α-subunits of the muscle receptor (i.e., at point "b," Fig. 7) leads to the β-subunit and a γ/ε/δ lineage (Kubo et al., 1985). That the γ- and ε-subunits diverged relatively recently is in line with the observation that, during muscle development, at least in mammals, a γ-subunit in the fetal muscle nACh receptor is replaced by an ε-subunit in the adult form (Witzemann et al., 1987).

The divergence of muscle and neuronal receptors is indicated by the separation of their α-subunits (i.e., *see* point "c," Fig. 7). This event is estimated to have occurred around 700–800 million years ago. On current evidence, this would have been rather early in metazoan evolution, and conceivably the branch point could mark the evolution of the developmental segregation of mesoderm and ectoderm. It is suggested that the original receptor hetero-oligomer was retained and evolved to become the muscle nACh receptor, and a new homo-oligomer was formed containing five of the neuronal-type α-subunits. That the neuronal/muscle divergence predates the separation of insects and vertebrates (point "d," Fig. 7) would suggest that the muscle of vertebrates and insects derived from a common origin: i.e., that these organisms (probably among others) evolved from common ancestry in which the muscle form of receptor was indeed already established. Paradoxically, glutamate and not acetylcholine is the excitatory neurotransmitter used at insect skeletal muscle (Ashford et al., 1987). This may reflect the poor selectivity of the ligand recognition site of the receptor or that different subunits evolved different types of recognition sites. In either case, subunits of the insect muscle glutamate receptor may be more similar to subunits of vertebrate muscle than they are to either vertebrate or invertebrate glutamate receptor subunits from neural tissue. Indeed, the pharmacology of the ion channel of the glutamate receptor in insect muscle shows similarity to that of vertebrate nACh receptors (Ashford et al., 1987,1988).

The first branch involving the neuronal α-subunit in the vertebrate lineage gives rise to the neuronal subunit, β2, and represents a second hetero-oligomerization event in the evolution of nACh receptors (point "e," Fig. 7). This event is estimated to have taken place around 600–700 million years ago. Since the β2-subunit is only distantly related to the β-subunit of the muscle receptor, it is surprising that this subunit substitutes for the β-subunit of the muscle receptor, but not for any of the other muscle receptor subunits, in functional expression studies (Deneris et al., 1988) It is possible that specific intersubunit contacts have been preserved in the β2-subunit positioned as it is between two α-subunits, allowing formation of the hybrid muscle receptor containing this subunit.

It is of interest to assess whether formation of the independent subtypes of the α-subunits of neuronal nACh receptor marks stages of expansion of the vertebrate nervous system. Almost certainly, the divergence of the α3-subtype from the branch leading to the α2- and α4-subtypes appears to have taken place early on in vertebrate evolution. This divergence may represent the formation of two distinct neuronal receptor types, one predominantly involved in autonomic control (Schoepfer et al., 1989) and the other involved in motor control. The divergence of branches leading to the α2 and α4-subtypes is estimated to have occurred around 300–400 million years ago and is the most recent duplication event in the evolution of neuronal nACh receptors shown on the tree. The α4-subtype is expressed at high levels throughout several distinct regions of the CNS, whereas the α2 is more restricted in its distribution (Wada et al., 1989; Daubas et al., 1990). This is an indication that the ancestral gene at this stage was probably of the α4-subtype.

Events In $GABA_A$ and Glycine Receptor Evolution

In the subtree of the receptor anion channels, the specialization of the three subtypes of α-subunit of the bovine $GABA_A$ receptor occurs much later in the tree than the separation of the $GABA_A$ and glycine receptors. Both of the duplication events involved in the formation of the α3-subtype and the α1- and α2-subtypes are considered to have taken place during the evolution of vertebrates (Hebebrand et al., 1987). Since the GABA-Bz receptor complex has also been identified in insects, such high evolutionary conservation suggests that there may indeed be an endogenous ligand for the Bz site (Robinson et al., 1986).

Since GABA and glycine are structurally similar, their receptors may have evolved from a proto-form that had weak selectivity in its recognition site. It is interesting to inquire what ligand was used by the ancestral form deduced from the tree. Probably it was glycine rather than GABA, since glycine is present in all cells, whereas GABA is not. Moreover, the evolution of the $GABA_A$ receptor must have depended on the evolution of glutamate decarboxylase (Jackson et al., 1990), which is required for GABA synthesis. However, once evolved, the restricted use of GABA as a chemical transmitter would have conferred a greater degree of specificity in signaling. A similar argument could be considered also for the excitatory transmitter receptors, since acetylcholine requires choline-acetyltransferase for its synthesis. Interestingly, for the anion channels, a glycine receptor has not yet been found in any invertebrate studied (Gerschenfeld, 1973), although the relevant studies on this point are perhaps as yet too few to exclude the possibility.

Concluding Remarks

Further understanding of the structural basis of the exacting physiological roles for diverse LGIC forms will be aided by a comparative approach to receptor studies. The elucidation of the structure of any one of the members of the LGIC superfamily will pave the way to a complete understanding of the distinct pharmacologies of LGICs, which ultimately requires an intimate and quantifiable knowledge of the 3D receptor structure.

The change in perspective brought about by the realization of the existence of the LGIC superfamily has itself created many new and interesting openings for further research. With the sequences at hand, those interested in receptors have started to direct studies to the dissection of structure and function using classical protein chemical and biochemical (Galzi et al., 1990; Abramson et al., 1989; Gomez et al., 1989) analyses, as well as site-directed mutagenesis, to create altered receptor forms. A variety of heterologous expression systems have been used to study the function of such receptors (Sumikawa et al., 1981; Pritchett et al., 1988; Paulson and Claudio, 1990). The possibility of having receptors from diverse sources in the same cellular background, and in isolation from the rest of the nervous system, should allow their pharmacologies to be directly compared and more precisely defined and quantified. Cell biologists have started to construct systems (Paulson and Claudio, 1990; Claudio et al., 1989) to examine aspects of receptor biology, such as how they are transported (Merlie, 1984) and regulated (Witzemann et al., 1987; Stollberg and Berg, 1987) and the physiological consequences of modifications at the primary structure level. Neuroanatomists have begun to use the nucleic acid probes and monospecific antibodies as molecular markers to delineate neuronal connections and circuitry and to classify neuronal cell and tissue types (Wada et al., 1989; Hebebrand et al., 1987; Wisdon et al., 1988; Montpied et al., 1988; Siegel, 1988). Molecular geneticists have started to identify the *cis*-elements (Klarsfeld et al., 1987; Baldwin and Burden, 1989; Berman et al., 1990) and *trans*-acting factors (Piette et al., 1989,1990) involved in controlling the expression of receptor genes, as well as mapping out their chromosomal location (Heidmann et al., 1986; Bessis et al., 1990; Grenningloh et al., 1990; Buckle et al., 1989) and organization (Nef et al., 1984; Boulter et al., 1990). Also, those interested in evolution, by using the relationships derived from analyses of receptor sequences, are beginning to gather information at the molecular level that may give insights into the ways in which nervous systems have evolved and into the factors that have influenced that evolution.

Acknowledgments

We gratefully acknowledge the support provided to the Molecular Modeling Unit at the University of Bath by R. J. Reynolds Tobacco Co; Shell Research and the SERC. We are most grateful to Dr. Susan Wonnacott for her many contributions to several aspects of this work and for her useful criticisms of the manuscript.

References

Abramson S. N., Culver P., Klines T., Li Y., Guest P., Gutman L., and Taylor P. (1988) Lophotoxin and related coral toxins covalently label the α-subunit of the nicotinic acetylcholine receptor. *J. Biol. Chem.* **263**, 18568–18573.

Abramson S. N., Li Y., Cuher P., and Taylor P. (1989) An analog of lophotoxin reacts covalently with Tyr-190 in the α subunit of the nicotinic ateylcholine receptor. *J. Biol. Chem.* **264**, 12666–12672.

Akagi H. and Miledi R. (1988) Heterogeneity of glycine receptors and their messenger RNAs in rat brain spinal cord. *Science* **242**, 270–273.

Albuquerque E. X., Tsai M-C., Aronstam R. S., Eldefrawi A. T., and Eldefrawi A. T. (1980) Sites of action of phencylclidine. *Mol. Pharmacol.* **18**, 167–178.

Aprison M. H. and Lipkowitz K. B. (1989) On the GABAa receptor: A molecular modelling approach. *J. Neuro. Sci.* **23**, 129–135.

Ashford M. L. J., Boden P., and Ramsey R. L. (1988) Voltage-dependent block of locust muscle glutamate channels by chlorisondamine. *J. Exp. Biol.* **50**, 745.

Ashford M. L. J., Boden P., Ramsey R. L., Shinozaki H., and Usherwood P. N. R. (1987) Effects of trimetaphan on locust muscle glutamate receptors. *J. Exp. Biol.* **130**, 405–424.

Atassi M. Z., Mulac-Jericevic B., Yokoi T., and Manshouri T. (1987) Localization of the functional sites on the α chain of the acetylcholine receptor. *FASEB* **46**, 2538–2547.

Bachmair A., Finley D., and Varshavsky A. (1986) In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234,** 179–186.

Baldwin T. J. and Burden S. J. (1989) Isolation and characterisation of the mouse acetylcholine receptor δ subunit gene: Identification of a 148-bp cis- acting region that confers myotube specific expression. *J. Cell. Biol.* **107,** 2271–2279.

Baldwin T. J., Yoshihara C. M., Blackner K., Kinter C. R., and Burden S. J. (1988) Regulation of acetylcholine receptor transcript expression during development in *Xenopus laevis*. *J. Cell Biol.* **106,** 469–478.

Barnard E. A., Darlison M. G., and Seeburg P. (1987) Molecular biology of the $GABA_A$ receptor: The receptor/channel superfamily. *Trends Neurosci.* **10,** 502–509.

Beeson D., Morris A., Vincent A., and Newsom-Davis J. (1990) The human muscle nicotinic acetylcholine receptor α-subunit exists as two isoforms: A novel exon. *EMBO J.* **9,** 2101–2109.

Bellone M., Tang F., Milius R., and Conti-Tronconi B. M. (1989) The main immunogenic region of the nicotinic acetylcholine receptor: Identification of amino acid residues interacting with different antibodies. *J. Immunol.* **143,** 3568–3579.

Berman S. A., Bursztajn S., Bowen B., and Gilbert W. (1990) Localization of an acetylcholine receptor intron to the nuclear membrane. *Science* **247,** 212–214.

Bessis A., Simon-Chazottes D., Devillers-Thiery A., Guenet J-L., and Changeux J-P. (1990) Chromosomal localization of the mouse genes coding for α2, α3, and α4 and β2 subunits of neuronal nicotinic acetylcholine receptor. *FEBS Lett.* **264,** 48–52.

Bishop M. J. and Friday A. E. (1985) Evolutionary trees from nucleic acid and protein sequences. *Proc. R. Soc. Lond.* **B226,** 271–302.

Blair L. A. C., Levitan E. S., Marshall J., Dionne V. E., and Barnard E. A. (1988) Single subunits of the $GABA_A$ receptor form ion channels with properties of the native receptor. *Science* **242,** 577–581.

Blount P. and Merlie J. P. (1988) Native folding of an acetylcholine receptor α subunit expressed in the absence of other receptor subunits. *J. Biol. Chem.* **263,** 1072–357.

Bormann J., Hamill O. P., and Sakmann B. (1987) Mechanism of anion permeation through channels gated by glycine and γ-aminobutyric acid in mouse cultured spinal neurones. *J. Physiol.* **385,** 243–286.

Bossy B., Ballivet M., and Spierer P. (1988) Conservation of neural nicotinic acetylcholine receptors from *Drosophila* to vertebrate central nervous systems. *EMBO J.* **7,** 611–618.

Boulanger T., Vercauteren D. P., Evrard G., and Durant F. (1989) Crystal structure and quantum electronic analyses of pitrazepin, an γ aminobutyric acid (GABA) receptor antagonist. *J. Chem. Soc. Perkin Trans. II* **3,** 217–221.

Boulter J., Connolly J., Deneris E., Goldman D., Heinemann S., and Patrick J. (1988) Functional expression of two neuronal nicotinic acetylcholine receptors from cDNA clones identifies a gene family. *Proc. Natl. Acad. Sci. USA* **84,** 7763–7767.

Boulter J., Evans K., Goldman D., Martin G., Treco D., Heinemann S., and Patrick J. (1986) Isolation of a cDNA clone coding for a possible neural nicotinic acetylcholine receptor α-subunit. *Nature* **31,** 368–374.

Boulter J., Luyten W., Evans K., Mason P., Ballivet M., Goldman D., Stengelin S., Martin G., Heinemann S., and Patrick J. (1985) Isolation of a clone coding for the α-subunit of a mouse acetylcholine receptor. *J. Neurosci.* **5,** 2545–2552.

Boulter J., OShea-Greenfield A., Duvoisin R. M., Connolly J. G., Wada E., Jensen A., Gardner P. D., Ballivet M., Deneris E. S., McKinnon D., Heinemann S., and Patrick J. (1990) Alpha3, alpha5, and beta4: Three members of the rat neuronal nicotinic acetylcholine receptor-related gene family form a gene cluster. *J. Biol. Chem.* **265,** 4472–4482.

Buckle V. J., Fujita N., Ryder-Cook A. S., Derry J. M. J., Barnard P. J., Lebo R. V., Schofield P. R., Seeburg P. H., Bateson A. N., Darlison M. G., and Barnard E. A. (1989) Chromosomal localization of $GABA_A$ receptor subunit genes: Relationship to human genetic disease. *Neuron* **3,** 647–654.

Buonanno A., Mudd J., and Merlie J. P. (1989) Isolation and characterisation of the β1 and ε-subunit genes of mouse muscle acetylcholine receptor. *J. Biol. Chem.* **264,** 7611–7616.

Burden S. J., DePalma R. L., and Gottesman G. S. (1982) Cross-linking of proteins in acetylcholine receptor-rich membranes: Association between the β-subunit and the 43 kD subsynaptic protein. *Cell* **35,** 687–692.

Bylund D. B. (1988) Subtypes of α_2-adrenoceptors: Pharmacological and molecular biological evidence converge. *Trends Pharmacol. Sci.* **9,** 356–361.

Carpenter G. and Zendegui J. G. (1986) Epidermal growth factor, its receptor, and related proteins. *Exp. Cell Res.* **104**, 1–10.

Carr C., McCourt D., and Cohen J. B. (1987) The 43-kilodalton protein of *Torpedo* nicotinic postsynaptic membranes: Purification and determination of primary structure. *Biochemistry* **26**, 7090–7102.

Carr C., Tyler A. N., and Cohen J. B. (1989) Myristic acid is the NH2-terminal blocking group of the 43-kDa protein of *Torpedo* nicotinic post-synaptic membranes. *FEBS Lett.* **43**, 65.

Cartaud J., Sobel A., Rousselet A., Devaux P. F., and Changeux J-P. (1981) Consequences of alkaline treatment for the ultra-structure of the acetylcholine receptor-rich membranes from *Torpedo marmaorata* electric organ. *J. Cell Biol.* **90**, 418–426.

Catterall W. A. (1988) Structure and function of voltage-sensitive ion-channels. *Science* **242**, 242–245.

Cauley K., Arganoff B. W., and Goldman D. (1989) Identification of a novel nicotinic acetylcholine receptor structural subunit expressed in goldfish retina. *J. Cell. Biol.* **108**, 636–645.

Chang H. W. and Neumann E. (1976) Dynamic properties of isolated acetylcholine receptor proteins release of calcium ions caused by acetylcholine binding. *Proc. Natl. Acad. Sci. USA* **73**, 3364–3368.

Charnet P., Labarca C., Leonard R. J., Vogelaar N. J., Czyzyk L., Gouin A., Davidson N., and Lester H. (1990) An open-channel blocker interacts with adjacent turns of α-helices in the nicotinic acetylcholine receptor. *Neuron* **2**, 87–95.

Chatrenet B., Tremeau 0., Botems F., Goeldner M. P., Hirth C. G. and Menez A. (1990) Topography of toxin-acetylcholine receptor complexes by using photoactivable toxin derivatives. *Proc. Natl. Acad. Sci. USA* **87**, 3378–3382.

Chen C-J., Chin J. E., Ueda K., Clarke D. P., Pastan I., Gonesman M. M., and Robinson I. B. (1986) Internal duplication and homology with bacterial transport proteins in the *mdr*1 (*P*-glycoprotein) gene from multidrug-resistant human cells. *Cell* **47**, 381–378.

Chiappinelli V. A., Wolf K. M., Grant G. A., and Chen S-J. (1990) κ_2-Bungarotoxin and κ_3-bungarotoxin: Two new neuronal nicotinic receptor antagonists isolated from the venom of Bungarus multicinctus. *Brain Res.* **509**, 237–248.

Chung F-Z. and Venter C. (1986) Molecular structure and evolution of adrenergic and cholinergic receptors. Proteins: Struct. Func. *Genet.* **1**, 287–301.

Claudio T., Ballivet M., Patrick J., and Heinemann S. (1983) Nucleotide and deduced amino acid sequence of *Torpedo californica* acetylcholine recep-tor γ–subunit. *Proc. Natl. Acad. Sci. USA* **80**, 1111–1115.

Claudio T., Paulson H. L., Green W. N., Ross A. F., Hartman D. S., and Hayden D. (1989) Fibroblasts transfected with *Torpedo* acetylcholine receptor β, γ, and δ subunit cDNAs express functional receptors when infected with a retroviral α recombinant. *J. Cell. Biol.* **108**, 2277–2290.

Cockcroft V. B., Osguthorpe D. J., Barnard E. A., and Lunt G. G. (1990) Modelling of agonist binding to the ligand-gated ion channel superfamily of receptors. *Proteins* (in press).

Conti-Tronconi B. M. and Raftery M. A. (1986) Nicotinic acetylcholine receptor contains multiple binding sites: Evidence form binding of α-dendrotoxin. *Proc. Natl. Acad. Sci. USA* **83**, 6646–6650.

Criado M., Witzemann V., Koenen M, and Sakmann B. (1988) Nucleotide sequence of the rat muscle acetylcholine receptor ε-subunit. *Nucl. Acids Res.* **16**, 10920.

Cull-Candy S. and Usowicz M. M. (1987) Patch-clamp recording from single glutamate receptor channels. *Trends in Pharmacol.* **8**, 218–224.

Daubas P., Devillers-Thiery A., Geofffroy B., Martinez S., Bessis A., and Changeux J-P. (1990) Differential expression of the neuronal acetylcholine receptor α2 subunit gene during chick brain development. *Neuron* **5**, 49–60.

Deisenhofer J., Epp O., Miki K., Huber R., and Michel H. (1985) Structure of the protein subunits in the photosynthetic reaction center of Rhodopseudomonas viridis at Å resolution. *Nature* **318**, 618–624.

Deneris E., Boulter J., Swanson L. W., Patrick J., and Heinemann S. (1989) β_3: A new member of nicotinic acetylcholine receptor gene family is expressed in brain. *J. Biol. Chem.* **264**, 6268–6272.

Deneris E. S., Connolly J., Boulter J., Wada E., Wada K., Swanson L. W., Patrick J., and Heinemann S. (1988) Primary structure and expression of β2: A novel subunit of neuronal acetvlcholine receptors. *Neuron* **1**, 45–54.

Dennis M., Giraudat J., Kotsyba-Hibert F., Goeldner M., Hirth C., Chang J., Lazure C., Chretien M., and Changeux J. (1988) Amino acids of the *Torpedo marmorata* acetylcholine receptor α subunit labeled by a photoaffinity ligand for the acetylcholine binding site. *Biochemistry* **27**, 2346–2357.

Derkach V., Surprenant A., and North R. A. (1989) 5-HT3 receptors are membrane ion channels. *Nature* **339**, 706–709.

DiPaola M., Czajkowski C., and Karlin A. (1989) The sidedness of the COOH terminus of the acetylcholine receptor subunit. *J. Biol. Chem.* **264**, 15457–15463.

Duggan M. J. and Stephenson F. A. (1989) Bovine γ-aminobutyric acid-A receptor sequencespecific antibodies: Identification of two epitopes which are recognised in both native and denatured γ-aminobutyric acid-A receptors. *J. Neurochem.* **53**, 132–139.

Dunn S. M. J., Conti-Tronconi B. M., and Raftery M. A. (1986) Acetylcholine receptor dimers are stabilised by extracellular disulphide bonding. *Biochemistry Biophys. Res. Comm.* **139**, 830–837.

Duvoisin R. M., Deneris E. S., Patrick J., and Heinemann S. (1989) The functional diversity of the neuronal nicotinic acetylcholine receptors is increased by a novel subunit: β4. *Neuron* **3**, 487–496.

Dwyer B. P. (1988) Evidence for the extramembranous location of the putative amphipathic helix of acetylcholine receptor. *Biochemistry* **27**, 5586–5592.

Finer-Moore J. and Stroud R. M. (1984) Amphipathic analysis and possible formation of the ion channel in an acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **81**, 155–159.

Frail D. E., L. L. McLaughlin, Mudd J., and Merlie J. P. (1988) Identification of the mouse muscle 43,000-daHon acetylcholine receptor associated protein (RAPsyn) by cDNA cloning. *J. Biol. Chem.* **263**, 15602–15607.

Furuichi T., Yoshikawa S., Miyawaki A., Wada K., Maeda N., and Mikoshiba K. Primary structure and functional expression of the inositol 1,4,5-trisphosphate-binding protein P_{400}. *Nature* **342**, 32–38.

Galligan J. J. and North R. A. (1990) MK-801 blocks nicotinic depoarizations of guinea pig myenteric neurons. *Neurosci. Lett.* **108**, 105–109.

Galzi J.-L., Revah F., Black D., Goeldner M., Hirth C., and Changeux J-P. (1990) Identification of a novel amino acid α-tyrosine 93 within the cholinergic ligand binding sites of the acetylcholine receptor by photoaffinity labeling. *J. Biol. Chem.* **265**, 10430–10437.

Geever R. F., Huiet L., Baum J. A., Tyler B. M., Patel V. B., Rutledge B. J., Case M. E., and Giles N. H. J. (1989) DNA sequence, organization and regulation of the qa gene cluster of Neurospora crassa. *J. Mol. Biol.* **207**, 15–34.

Gerschenfeld H. M. (1973) Chemical transmission in invertebate central nervous systems and neuromuscular junctions. *Physiol. Rev.* **53**, 1–117.

Giraudat J., Dennis M., Heidmann T., Haumont P-Y., Lederer F., and Changeux J-P. Structure of the high-affinity binding site for noncompetitive blockers of the acetylcholine receptor: [^{3}H]chlorpromazine labels homologous residues in the β and δ chains. *Biochemistry* **26**, 2410–2418.

Godzik A. and Sander C. (1989) Conservation of residue interactions in a family of Ca-binding proteins. *Prot. Engineer.* **2**, 589–596.

Gomez A. R., Femandez-Shaw C., Valdivleso F., and Mayor F. (1989) Chemical modification of the glycine receptor with fluorescein isothiocyanate specifically affects the interaction of glycine with its binding site. *Biochemistry Biophys. Res. Commun.* **160**, 374–381.

Gregor P., Mano I., Maoz I., McKeown M., and Teichberg V. I. (1989) Molecular structure of the chick cerebellar kainate-binding subunit of a putative glutamate receptor. *Nature* **342**, 689–692.

Grenningloh G., Rienitz A., Schmin B., Methfessel C., Beyrether K., Gundelfinger E. D., and Betz H. (1987) The strychnine binding subunit of the glycine receptor shows homology with the nicotinic acetylcholine receptor. *Nature* **328**, 215–220.

Grenningloh G., Schmieden V., Schofield P. R., Seeburg P. H., Siddique T., Mohandas T. K., Becker C., and Betz H. (1990) Alpha subunit variants of the human glycine receptor: Primary structures, functional expression and chromosomal localization of the corresponding genes. *EMBO J.* **9**, 771–776.

Griesmann G E., McCormick D. J., Aizpurua H. J., and Lennon V. A. (1990) α-bungarotoxin binds to human acetylcholine receptor alpha-subunit peptide 185–199 in solution but not to peptides 125–147 and 389–409. *J. Neurochem.* **54**, 1541–1547.

Hardie R. C. (1989) A histamine-activated chloride channel in neurotransmission at a photoreceptor synapse. *Nature* **339**, 704–706.

Hebebrand J., Friedl W., Kessler K., Hewing M., and Propping P. (1987) Further characterisation of the avian benzodiazepine receptor subunits including phylo- and ontogenetic aspects. *J. Neurochem.* **49**, 1888–1893.

Heidmann O., Buonanno A., Geoffroy B., Robert B., Guenet J-L., Merlie J. P., and Changeux J-P. (1986) Chromosomal localisation of muscle nicotinic acetylcholine receptor genes in the mouse. *Science* **234**, 866–868.

Henderson, Komiya H., Yeates T. O., Rees D. C., Allen J. P., and Feher G. (1988) Structure of the reaction center from *Rodobacter sphaeroides* R-26 and 2.4.1: Symmetry relations and sequence comparisons between different species. *Proc. Natl. Acad. Sci. USA* **85**, 9012–9016.

Higgins C. (1989) Protein joins transport family. *Nature* **340**, 342.

Hollman M. O., Shea-Greenfield A., Rogers S. W., and Heinemann S. (1989) Cloning by functional expression of a member of the glutamate receptor family. *Nature* **342**, 643–648.

Hucho F., Oberthur W., and Lottspeich F. (1986) The ion channel of the nicotinic acetylcholine receptor is formed by the homologous helices MII of the receptor subunits. *FEBS Lett.* **205**, 137–142.

Huettner J. E. and Bean B. P. (1988) Block of *N*-methyl-D-aspartate-activated current by the anticonvulsant MK-801: Selective binding to open channels. *Proc. Natl. Acad. Sci. USA* **85**, 130–1311.

Imoto K., Methfessel C., Sakmann B., Mishina M., Mori Y., Konno T., ukuda K., Kurasaki M., Bujo H., Fujita Y., and Numa S. (1986) Location of a α-subunit region determining ion transport through the acetylcholine receptor channel. *Nature* **324**, 670–674.

Imoto K., Busch C., Sakmann B., Mishina M., Konno T., Nakai J., Bujo H., Mori Y., Fukuda K., and Numa S. (1988) Rings of negatively charged amino acids determine the acetylcholine receptor channel conductance. *Nature* **335**, 645–648.

Jackson F. R., Newby L. M., and Kulkarni S. J. (1990) *Drosophila* GABAergic systems: Sequence and expression of glutamic acid decarboxylase. *J. Neurochem.* **54**, 1068–1078.

Jackson T. R., Blair L. A. C., Marshall J., Goedert M., and Hanley M. R. (1988) The *mas* oncogene encodes an angiotensin receptor. *Nature* **335**, 437–440.

Jan L. Y. and Jan Y. N. (1990) A superfamily of ion channels. *Nature* **345**, 672.

Jing S., Spencer T., Miller K., Hopkins C., and Trowbridge I. S. (1990) Role of the human transferrin receptor cytoplasmic domain in endocytosis: Localisation of a specific signal sequence for internalisation. *J. Cell Biol.* **110**, 283–294.

Johnson D. A., Cushman R., and Malekzadeh R. (1990) Orientation of cobra α-toxin on the nicotinic acetylcholine receptor. *J. Biol. Chem.* **265**, 7360–7368.

Kanne D. B. and Abood L. G. (1988) Synthesis and biological characterization of pyridohomotropanes. Structure-activity relationships of conformationally restricted nicotinoids. *J. Med. Chem.* **31**, 506–509.

Karlin A. (1980) Molecular properties of nicotinic acetylcholine receptors. *The Cell Surface and Neuronal Function.* Poste G., ed., pp. 191–260.

Kaupp U. B., Niidome T., Tanabe T., Terada T., Bonigk S., Stuhmer W., Cook W., Kangawa N., Matsu K., Hirose H., and Miyata T. (1989) Primary structure and functional expression from complementary DNA of the rod photoreceptor cyclic GMP-gated channel. *Nature* **342**, 762–766.

Kavanaugh M. P., Tester B. A. C., and Weber E. (1989) Interaction of MK-801 with the nicotinic acetylcholine receptor-associated ion channel from electroplax. *Eur. J. Pharmacol.* **164**, 397–398.

Klarsfeld A., Daubas P., Bourachot B., and Changeux J.-P. (1987) A 5'-flanking region of the chicken acetylcholine receptor α-subunit gene confers tissue specificity and developmental control of expression in transfected cells. *Mol. Cell. Biol.* **7**, 951–955.

Kosower E. M. (1988) A partial structure for the γ-aminobutyric acid ($GABA_A$) receptor is derived from the model for the nicotinic acetylcholine receptor: The anion-exchange protein of cell membranes is related to the GABAa receptor. *FEBS Lett.* **231**, 5–40.

Krogsgaard-Larsen P., Mikkelsen H., Jacobsen P., Falch E., Curtis D. R., Peet M. J., and Leah J. D. (1983) 4,5,6,7-Tetrahydroisoxazolo[5,4-c]pyridin-3-ol and related analogues of THIP. Synthesis and biological activity. *J. Med. Chem.* **26**, 895–900.

Kubalek E., Ralston S., Lindstrom J., and Unwin N. (1987) Location of subunits within the acetylcholine receptor by electron image anaysis of tubular crystals from *Torpedo marmorata.* **105**, 9–18.

Kubo T., Noda M., Takai T., Tanabe T., Kayano T., Shimizu S., Tanaka K., Takahashi H., Hirose T., Inayama S., Kikuno R., Miyata T., and Numa S. (1985) Primary structure of δ-subunit precursor of calf muscle acetylcholine receptor deduced from cDNA sequence. *Eur. J. Biochem.* **149**, 5–13.

Kurosaki T., Fukuda K., Mori Y., Tanaka K., M. Mishina and Numa S. (1987) Functional properties of nicotinic acetylcholine receptor subunits expressed in various combinations. *FEBS Lett.* **2**, 253–258.

Lambolez B. and Rossier J. (1987) Benzodiazepine agonists protect a histidine residue from modification by diethyl pyrocarbonate whereas propyl β-carboline does not. *FEBS Lett.* **219**, 301–305.

Langenbuch-Cachat J., Bon C., Mulle C., Goeldner M., Hirth C., and Changeux J. (1988) Photoaffinity labeling of the acetylcholine binding on the nicotinic receptor by an aryldiazonium derivative. *Biochemistry* **27**, 2337–2345.

Langosch D., Thomas L., and Betz H. (1988) Conserved quaternary structure of ligand-gated ion channels: The postsynaptic glycine receptor is a pentamer. *Proc. Natl. Acad. Sci. USA* **85**, 7394–7398.

LaPolla R. J., Mayne K. M., and Davidson N. (1984) Isolation and characterisation of a cDNA clone for the complete protein coding region of the δ subunit of the mouse acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **81**, 7970–7974.

Lear J. D., Wasserman Z. R., and DeGrado W. F. (1988) Synthetic peptide models for protein ion channels. *Science* **240**, 1177–1181.

Lefkowitz R. J. and Caron M. G. (1988) Models for the study of receptors coupled to guanine nucleotide regulatory proteins. *J. Biol. Chem.* **263**, 4993–4996.

Leonard R. J., Labarca C. G., Charnet P., Davidson N., and Lester H. A. (1988) Evidence that the M2 membrane-spanning region lines the ion channel pore of the nicotinic acetylcholine receptor. *Science* **242**, 1578–1581.

Levitan E. S., Schofield P. R., Burt D. R., Rhee L. M., Wisden W., Kohler M., Fujita N., Rodriguez H., Stephenson A., Darlison M. G., Barnard E. A., and Seeburg P. H. (1988) Structural and functional basis for $GABA_A$ receptor heterogeneity. *Nature* **335**, 76–79.

Lindstrom J., Schoepfer R., and Whiting P. (1987) Molecular studies of the neuronal nicotinic acetylcholine receptor family. *Prog. Mol. Neurobiol.* **1**, 281–327.

Lo M. S., Garland P. B., Lamprecht J., and Barnard E. A. (1980) Rotational mobility of the membrane-bound acetylcholine receptor of *Torpedo* electric organ measured by phosphorescence depolarisation. *FEBS Lett.* **111**, 407–412.

Luetje C. W., Wada K., Rogers S., Abramson S. N., Tsuji K., Heinemann S., and Patrick J. (1990) Neurotoxins distinguish between different neuronal nicotinic acetylcholine receptors. *J. Neurochem.* **50**, 632–640.

Lukas R. L. (1989) Pharmacological distinctions between functional nicotinic acetylcholine receptors on the PC12 rat pheochrornocytoma and the TE671 human medulloblastoma. *J. Pharm. Exp. Ther.* **251**, 175–182.

Luyten W. H. M. L. (1986) A model for the acetylcholine binding site of the nicotinic acetylcholine receptor. *J. Neurosci. Res.* **16**, 51–73.

Maelicke A., Plumer-Wilk R., Fels G., Spencer S. R., Engelhard M., Veltel D., and Conti-Tronconi B. M. (1989) Epitope mapping employing antibodies raised against short synthetic peptides: A study of the nicotinic acetylcholine receptor. *Biochemistry* **28**, 1396–1405.

Maksay G. and Ticku M. K. (1984) Characterisation of γ-aminobutyric acid-benzodiazepine receptor complexes by protection against inactivation by group-specific reagents. *J. Neurochem.* **42**, 1715–1727.

Masu Y., Nakayama K., Tamaki H., Harada Y., Kuno M., and Nakanishi S. (1987) cDNA cloning of bovine substance-K receptor through oocyte expression system. *Nature* **329**, 836–838.

McLane K. E., Wu X., and Conti-Tronconi B. M. (1990a) Identification of a brain acetylcholine receptor subunit able to bind α-bungarotoxin. *J. Biol. Chem.* **265**, 9816–9824.

McLane K. E., Tang F., and Conti-Tronconi B. M. (1990b) Localisation of sequence segments forming a kappa-bungarotoxin-binding site on the alpha3 neuronal nicotinic receptor. *J. Biol. Chem.* **265**, 1537–1544.

Merlie J. P. (1984) Biogenesis of the acetylcholine receptor, a multi-subunit integral membrane protein. *Cell* **36**, 573–575.

Mielke D. L. and Wallace B. A. (1988) Secondary structural analyses of the nicotinic acetylcholine receptor as a test of molecular models. *J. Biol. Chem.* **263**, 3177–3182.

Mishina M., Tobimatsu T., Irnoto K., Tanaka K., Fulita Y., Fukuda Y., Hirose T., Inayama S., Takahashi T., Kuno M., and Numa S. (1985) Location of functional regions of acetylcholine receptor α-subunit by site-directed mutagenesis. *Nature* **313**, 364–369.

Mitra M., McCarthy M. P., and Stroud R. M. (1989) Three-dimensional structure of the nicotinic acetylcholine receptor and location of the major associated 43-kD cytoskeletal protein, determined at 22 Å by low dose electron micorscopy and X-ray diffraction to 12.5 Å. *J. Cell. Biol.* **109**, 755–774.

Montpied P., Martin B. M., Cottingham S. L., Stubblefield B. K., Ginnis. E. 1., and Paul S. M. (1988) Regional distribution of the $GABA_A$/benzodiazepine receptor mRNA in rat brain. *J. Neurochem.* **51**, 1651–1654.

Mulac-Jericevic B., Kurisaki J., and Atassi M. Z. (1987) Profile of the conlinuous antigenlc regions on the extracellular part of the α chain of an acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **84**, 3633–3637.

Musil L. S., Frail D. E., and Merlie J. P. (1989) The mammalian 43-kD acetylcholine receptor associated protein (RAPsyn) is expressed in some nonmuscle cells. *J. Cell. Biol.* **108**, 1833–1840.

Nef P., Mauron A., Stalder R., Alliod C., and Ballivet M. (1984) Structure, linkage, and sequence of the two genes encoding the δ and γ subunits of the nicotinic acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **81**, 7975–7979.

Nef P., Oneyser C., Alliod C., Couturier S., and Ballivet M. (1988) Genes expressed in the brain define three distinct neuronal nicotinic acetylcholine receptors. *EMBO J.* **7**, 595–601.

Neher, E. and Steinbach, J. (1978) Local anesthetics transiently block currents through single acetylcholine receptor channels. *J. Physiol.* **277**, 153–176.

Nelson H., Mandiyan S., and Nelson N. (1990) Cloning of the human brain GABA transporter. *FEBS Lett.* **269**, 181–184.

Neumann D., Barchan D., Horowitz M., Kochva E., and Fuchs S. (1989) Snake acetylcholine receptor: Cloning of the domain containing the four extracellular cysteines of the α subunit. *Proc. Natl. Acad. Sci. USA* **86**, 7255–7259.

Noda M., Ikeda T., Kayano T., Suzuki H., Takeshima H., Kurasaki M., Takahashi H., and Numa S. (1986) Existence of distinct sodium channel messenger RNAs in rat brain. *Nature* **320**, 188–192.

Noda M., Takahashi H., Tanabe T., Toyosato M., Furutani Y., Hirose T., Asai M., Inayama S., Miyata T., and Numa S. (1982) Primary structure of α-subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequence. *Nature* **299**, 793–797.

Noda M., Takahashi H., Tanabe T., Toyosato M., Kikyotani S., Furutani Y., Hirose T., Takashima H., Inayama S., Miyata T., and Numa S. (1983a) Structural homology of *Torpedo californica* acetylcholine receptor subunits. *Nature* **302**, 528–532.

Noda M., Furutani Y., Takahashi H., Toyosato M., Tanabe T., Shimizu S., Kikyotani S., Kayano T., Hirose T., Inayama S., and Numa S. (1983b) Cloning and sequence analysis of calf cDNA and human genomic DNA encoding α-subunit precursor of muscle acetylcholine receptor. *Nature* **305**, 818–823.

O'Dell T. L. and Christensen B. N. (1988) Mecamylamine is a selective non-competitive antagonist of *N*- methyl–D-aspartate and aspartate-induced currents in horizontal cells dissociated from the catfish retina. *Neurosci. Lett.* **94**, 93.

Paulson H. L. and Claudio T. (1990) Temperature-sensitive expression of all-*Torpedo* and *Torpedo*-rat hybrid AChR in mammalian muscle cells. *J. Cell Biol.* **110**, 1705–1717.

Pedersen S. E. and Cohen J. B. (1990)*d*-Tubocurarine binding sites are located at α-γ and α–δ subunit interfaces of the nicotinic acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* **87**, 2785–2789.

Persson B., Flinta C., Heijne G. voc, and Jornvall, H. (1985) Structures of *N*-terminally acetylated proteins. *Eur. J. Biochem.* **152**, 523–527.

Piette J., Bessereau J. L., Huchet M., and Changeux J-P. (1990) Two adjacent MyoD-1-binding sites regulate expression of the acetylcholine receptor α-subunit gene. *Nature* **345**, 353–355.

Piette J., Klarsfeld A., and Changeux J-P. (1989) Interaction of nuclear factors with the upstream region of the alpha-subunit gene of chicken muscle acetylcholine receptor: Variations with muscle differentiation and denervation. *EMBO J.* **8**, 687–694.

Pongs O., Kecskemethy N., Muller R., Krah-Jentgens I., Baumann A., Kiltz, H. H., Canal I., Llamazares S., and Femus A. (1988) Shaker encodes a family of putative potassium channel proteins in the nervous system of *Drosophila*. *EMBO J.* **7**, 1087–1095.

Pritchett D. B., Sontheimer H., Kettenmann C. M., Seeburg P. H., and Schofield P. R. (1988) Transient expression shows ligand gating and allosteric potentiation of $GABA_A$ receptor subunits. *Science* **242**, 1306–1308.

Pritchett D. B., Sontheimer H., Shivers B. D., Ymer S., Kettenmann H., Schofield P. R., and Seeburg P. H. (1989) Importance of a novel GABAA receptor subunit for benzodiazepine pharmacology. *Nature* **338**, 582–585.

Qu Z., Moritz E., and Huganir R. L. (1990) Regulation of tyrosine phosphorylation of the nicotinic acetylcholine receptor at the rat neuromuscular junction. *Neuron* **2**, 367–378.

Raftery M. A., Hunkapiller M. W., Strader C. D., and Hood L. E. (1980) Acetylcholine receptor: Complex of homologous subunits. *Science* 208, 1454–1457.

Ramoa A. S., Alkondon M., Aracava Y., Irons J., Lunt G. G., Deshphande S. S., Wonnacott S., and Albuquerque E. X. (1990) The anticonvulsant MK–801 blocks peripheral and central nicotinic acetylcholine receptor ion channels. *J. Pharmacol. Exp. Ther.* **254**, 71–82.

Rapier C., Lunt G. G., and Wonnacott S. (1990) Nicotinic modulation of [^{3}H]dopamine release from striatal synaptosomes: Pharmacological characterisation. *J. Neurochem.* **54**, 937–945.

Rees D. C., DeAntonio L., and Eisenberg D. (1989) Hydrophobic organisation af membrane proteins. *Science* **245**, 510–513.

Reynolds I. J. and Miller R. J. (1988) Tricyclic antidepressants block *N*-methyl–D–aspartate receptors: Similarities to the action of zinc. *Brit. J. Pharmacol.* **95**, 95–102.

Riordan J. A., Rommens J. M., Kerem B–S., Alon N., Rozmahel R., Grzelczak Z., Zielenski J., Lok S., Plavsic N., Chou J–L., Drumm M. L., Iannuzzi M. C., Collins F. S., and Tsui L–C. (1989) Identification of the cystic fibrosis gene: Cloning and characterisation of complementary DNA. *Science* **245**, 1066–1072.

Robinson R. N., Macallen D., Lunt G. G., and Weitzman P. D. J. (1986) The GABA receptor of insect CNS: Characterisation of a benzodiazepine binding site. *J. Neurochem.* **47**, 1955–1961.

Roth B., Schwendimann B., Hughes G. J., Tzartos S. J., and Barkas T. (1987) A modified nicotinic acetylcholine receptor lacking the "ion channel amphipathic helices." *Febs Lett.* **221**, 172–178.

Saedi M. S., Anand R., Conroy W. G., and Lindstrom J. (1990) Determination of amino acids critical to the main immunogenic region of intact acetylcholine receptors by in vitro mutagenesis. *FEBS Lett.* **267**, 55–59.

Saimi Y., Matinac B., Gustin M. C., Culbertson M. R., Alder J., and Kung C. (1988) Ion channels in *paramecium*, yeast and *Eschericia coli. TIBS* **8**, 304–309.

Sarkar H. B., Thorens F., Lodish H., and Kaback H. (1988) Expression of the human erythrocyte glucose transporter in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **85**, 5463–5467.

Schmieden V., Grenningloh G., Schofield P. R., and Betz H. (1989) Functional expression in *Xenopus* oocytes of the strychnine binding 48 kD subunit of the glycine receptor. *EMBO J.* **8**, 695–699.

Schoepfer R., Halvorsen S W., Conroy W. G., Whiting P., and Lindstrom J. (1989) Antisera against an acetylcholine receptor α3 fusion protein bind to ganglionic but not to brain nicotinic acetylcholine receptors. *FEBS Lett.* **257**, 393–399.

Schofield P. R., Darlison M. G., Fujita N., Burt D., Stephenson F. A., Rodriguez H., Rhee L. M., Ramachandran J., Reale V., Glencorse T. A., Seeburg P. H., and Barnard E. A. (1987) Sequence and functional expression of the GABAA receptor shows a ligand gated receptor superfamily. *Nature* **328**, 221–227.

Siegel R. E. (1988) The mRNAs encoding $GABA_A$/benzodiazepine receptor subunits are localized in different cell populations of the bovine cerebellum. *Neuron* **1**, 579–584.

Stephenson F. A. (1988) Understanding the GABA–A receptor: A chemically gated ion channel. *Biochemistry J.* **249**, 21–32.

Stollberg J. and Berg D. K. (1987) Neuronal acetylcholine receptors: Fate of surface and internal pools in cell culture. *J. Neurosci.* **7**, 1809–1815.

Sumikawa K., Houghton M., Emtage J. S., Richards B. M., and Barnard E. A. (1981) Active multi–subunit ACh receptor assembled by translation of heterologous mRNA in Xenopus oocytes. *Nature* **292**, 862–864.

Takai T., Noda M., Mishina M., Shimizu S., Furutani Y., Kayano T., Ikeda T., Kubo T., Takahashi H., Takahashi T., Kuno M., and Numa S. (1985) Cloning, sequencing and expression of cDNA for a novel subunit of acetylcholine receptor from calf muscle. *Nature* **315**, 761–764.

Takeshima H., Nishimura S., Matsumoto T., Ishida H., Kangawa K., Minamino N., Matsuo H., Ueda M., Hanaoka M., Hirose T., and Numa S. (1989) Primary structure and expression from complementary DNA of skeletal muscle ryanodine receptor. *Nature* **339**, 439–445.

Tanabe T., Takeshima H., Mikami A., Flockerzi V., Takahashi H., Kangawa K., Kojima M., Matsuo H., Hirose T., and Numa S. (1987) Primary structure of the receptor for calcium channel blockers from skeletal muscle. *Nature* **328**, 313–318.

Tobimatsu T., Fujita Y., Fukuda K., Tanaka K., Mori Y., Konno T., Mishina M., and Numa S. (1987) Effects of substitution of putative transmembrane segments on nicotinic acetylcholine receptor function. *FEBS Lett.* **222**, 56–62.

Toyoshima C. and Unwin N. (1988) Ion channel of acetylcholine receptor reconstructed from images of postsynaptic membranes. *Nature* **336**, 247–250.

Tzartos S. J., Loutrari H. V., Tang F., Kokla A., Walgrave S. L., Milius R. P., and Conti-Troconi B. M.

(1990) Main immunogenic region of *Torpedo* electroplax and human muscle acetylcholine receptor: Localization and microheterogeneity revealed by the use of synthetic peptides. *J. Neurochem.* **54,** 51–61.

Wada A., Uezono Y., Arita M., Tsuji K., Yanagihara K., Kobayashi H., and Izumi F. (1989) High-affinity and selectivity of neosurugatoxin for the inhibition of ^{22}Na influx via nicotinic receptor–ion channel in cultured bovine adrenal medullary cells: Comparative study with histrionicotoxin. *Neurosci.* **33,** 333.

Wada E., Wada K., Boulter J., Deneris E. S. Heinemann, Patrick J., and Swanson L. W. (1989) Distribution of alpha2, alpha3, alpha4, and beta2 neuronal nicotinic receptor subunit mRNAs in the central nervous system: A hybridization histochemical study in the rat. *J. Comp. Neurol.* **284,** 314–335.

Wada K., Ballivet M., Boulter J., Connolly J., Wada E., Deneris E. S., Swanson L. W., and Heinemann S. (1988) Functional expression of a new pharmacological subtype of brain nicotinic acetylcholine receptor. *Science* **240,** 330–334.

Wada K., Dechesne C. J., Shimasaki S., King R. G., Kusano K., Buonanno A., Hampson D. R., Banner C., Wenthold R. J., and Nakatani Y. (1989) Sequence and expression of a frog brain complementary DNA encoding a kainate–binding protein. *Nature* **342,** 684–689.

Ward J., Cockcroft V. B., Smillie F., Lunt G., and Wonnacott S. (1990) Methyllycaconitine: A selective probe of neuronal bungarotoxin binding proteins. *FEBS Lett.* **270,** 45–48.

Widdows K. B., Kirkness E. F., and Tumer A. J. (1987) Modification of the GABA/benzodiazepine receptor with the arginine reagent, 2,3–butadione. *FEBS Lett.* **222,** 125–128.

Wilson P. T., Hawrot E., and Lentz T. L. (1988) Distribution of α-bungarotoxin binding sites over residues 173–204 of the α-subunit of the acetylcholine receptor. *Mol. Pharmacol.* **34,** 643–650.

Wisden W., Morris B. J., Darlison M. G., Hunt S. P., and Barnard E. A. (1988) Distinct $GABA_A$ receptor α subunit mRNAs show differential patterns of expression in bovine brain. *Neuron* **1,** 937–947.

Witzemann V., Barg B., Nishikawa Y., Sakmann B., and Numa S. (1987) Differential regulation of muscle acetylcholine receptor γ- and ε-subunit mRNA. *FEBS Lett.* **223,** 104–112.

Wonnacott S. (1987) Neurotoxin probes for neuronal nicotinic receptors. *Neurotoxins and Their Pharmacological Implications.* Jenner P., ed., Raven, New York, pp. 209–231.

Yakel J. L. and Meyer M. B. (1988) $5\text{-}HT^3$ receptors mediate rapid responses in cultured hippocampus and a clonal cell line. *Neuron* **1,** 615–621.

Yamada S., Kagawa Y., Takayanagi N., Nakayama K., Tsuji K., Kosuge T., Hayashi E., Okada K., and Inoue S. (1987) Comparison of antinicotinic activity by neosurugatoxin and the structurally related compounds. *J. Pharm. Expt. Ther.* **243,** 1153–1158.

Ymer S., Schofield P. R., Draguhn A., Werner P., Kohler M., and Seeburg P. H. (1989) $GABA_A$ receptor beta subunit heterogeneity: Functional expression of cloned cDNAs. *EMBO J.* **8,** 1665–1670.