Functional and Structural Divergence Prediction of Proteins from Molecular Phylogenetic Analysis with Special Reference to Energy Metabolism System and Nervous System

Takeshi Tomiki

Contents

1. General Introduction

Organisms are living by using combinations of various functional systems. How have functional systems been diversified through evolution? Each system is composed of various proteins, therefore functional systems have been diversified by changing functions of proteins in each system. How were protein functions in functional systems diversified? Amino acid sequences of each protein have been substituted through evolution, and protein functions of each protein have been changed. The main theme of this dissertation is to predict such functional changes through molecular phylogenetic analyses.

Electron transfer energy metabolism system is the one of the good targets for the analysis about the divergence of protein functions and functional systems. There are various electron transfer energy metabolism systems, but the common characteristic features are shared among the systems. For example among the four major electron transfer energy metabolism systems (photosynthesis, aerobic recpiration, denitrification and sulfur respiration), ATP synthase for producing ATP and quinone as an electron transporter are used in all the four systems, and cytochrome c is used as an electron transporter in photosynthesis, aerobic respiration and denitrification (Fig. 1.1). Furthermore, homologous proteins are found among photosynthesis, aerobic respiration and denitrification (Berry et al. 2000; Hauska et al. 1988), between aerobic respiration and denitrification (Mogi et al. 1998; Saraste and Castresana 1994; Zumft et al. 1992) and between denitrification and sulfur respiration (Dias et al. 1999; Krafft et al. 1992; Moreno-Vivian et al. 1999) (Fig. 1.1). These common characteristics and homologous proteins in the four systems suggest that there are evolutionary relationships among the four systems. But the previous studies about the homologous proteins are not enough to discuss how these four systems have evolved and how

the protein functions in the four systems have changed. Previous studies could suggest the phylogeny of each homologous protein in the four systems, but they could not suggest the phylogeny of the four systems. If we want to describe the phylogeny of the four systems, we should refer not only one phylogenetic analysis of each homologous protein but also multiple phylogenetic analyses of homologous proteins in the four systems. Protein functions have changed by amino acid substitutions through evolution. We should describe protein function changes referring to both phylogenetic trees of each homologous protein and amino acid sequences of each protein. To predict how functional systems have evolved and how protein functions have been changed through evolution, I conducted phylogenetic analyses at three levels; system, domain, amino acid.

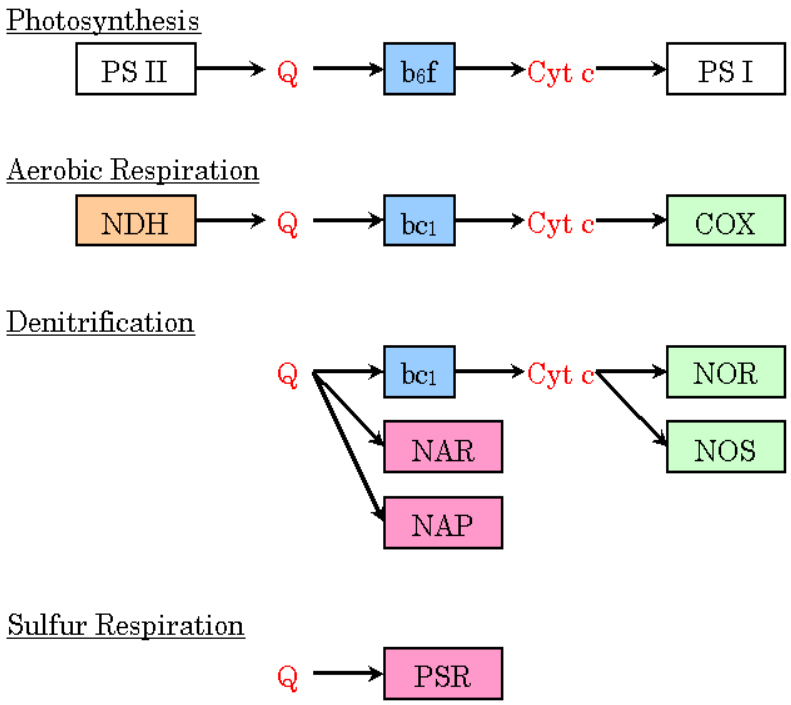


Fig. 1.1: Homologous proteins in photosynthesis, aerobic respiration, denitrification and sulfur respiration. Black arrows designate electron flows. Square objects are proteins. Same colored proteins are homologous. Abbreviations in this figure are as follows. PS II = photosystem II; Q = Quinone; b_6f = cytochrome b_6f complex; Cyt c= cytochrome c; PS I = photosystem I; NDH = NADH dehydrogenase; bc_1 = cytochrome bc_1 complex; COX = cytochrome c oxidase; NAR = respiratory nitrate reductase; NAP = periplasmic nitrate reductase; NOR = nitric oxide reductase; NOS = nitrous oxide reductase; PSR = polysulfide reductase

In this study, I tried to predict the evolution of functional systems by superimposition of phylogenetic trees constructed by as many homologous proteins in the functional systems as possible. The basic idea of the predicting method was based on the phylogenetic analysis of muscle tissues by OOta and Satitou (1999). There are eight muscle tissue classes in vertebrates and invertebrates. If homologous genes are expressed in different tissue classes, the divergences of structural genes in the eight classes correspond to the divergence of the tissue classes, In this case, the inference of deduced tissue phylogenetic trees is possible from phylogenetic trees of structural genes. If the same gene is expressed in more than one tissue classes, it can be interpreted that the tissue classes are closely related. In this case, phylogenetic trees help us to infer deduced tissue phylogenetic trees. Under the assumption that the five structural genes they chose are these two cases, OOta and Saitou (1999) superimposed phylogenetic trees of five structural genes in the eight muscle tissue classes, and inferred the deduced phylogenetic tree of muscle tissues.

I applied this idea to the inference of the deduced phylogenetic tree of functional systems. Under the assumption that phylogenetic trees constructed by structural genes in functional systems correspond to deduced phylogenetic tree of functional systems, I tried to superimpose phylogenetic trees of as many homologous proteins in the functional systems as possible to infer the deduced phylogenetic tree of functional systems. I did phylogenetic analysis at system level to predict phylogeny of functional systems.

Some of proteins are composed of multiple functional domains, and each domain might be inserted or deleted through evolution. Insertions or deletions of functional domains have changed protein functions. I tried to predict how domain compositions had changed through evolution and how protein functions had diversified by domain insertions and deletions. Phylogenetic trees constructed by amino acid sequences of proteins are not enough

to predict domain composition changes in some cases. In these cases, multiple phylogenetic trees constructed by amino acid sequences of various functional domains are necessary. I did phylogenetic analysis at the domain level by constructing composite gene trees so as to predict protein function changes in functional systems.

It is possible that protein functions are different even if domain compositions of proteins are the same. I tried to predict the essential amino acid substitutions for changing protein functions through evolution referring to phylogenetic trees and amino acid sequences which cannot be detected by domain level analysis. I did phylogenetic analysis at the amino acid level to predict protein function changes in functional systems.

As I described above, I predicted the evolution of functional systems and protein function changes in the functional systems at the three levels; system level, domain level and amino acid level. I applied the three level predictions to the two functional systems; electron transfer energy metabolism system (chapter 2) and neurotransmission system (chapter 3). As I explained in this section, the four electron transfer energy metabolism systems (photosynthesis, aerobic respiration, denitrification and sulfur respiration) share the common characteristic features (Fig. 1.1). Furthermore, homologous proteins exist among the four systems (Berry et al. 2000; Dias et al. 1999; Hauska et al. 1988; Krafft et al. 1992; Mogi et al. 1998; Moreno-Vivian et al. 1999; Saraste and Castresana 1994; Zumft et al. 1992) (Fig. 1.1). They suggest that the four systems are evolutionarily related. But the exhaustive phylogenetic analysis of proteins among the four systems has not been tried yet. Furthremore, I had done phylogenetic analyses of proteins in the four energy metabolisms in master course. Therefore, I decided to try the three level predictions in the four energy metabolism systems to infer the phylogeny of the four systems and the protein function changes in the four systems. I constructed molecular phylogenetic trees by using amino acid sequences of

functional domains. These trees and amino acid compositions of proteins suggest that domain insertions and deletions in the four systems made functions of electron transfer in proteins change. I tried to predict ligand binging specificities of catalytic proteins at the amino acid level, but experimental data about ligand binding functions are not enough for doing such predictions. Therefore, I did not predict protein function changes at amino acid level. Most of proteins in the four systems are not homologous each other. Only some important proteins for generating energy are homologous among the four systems. It means that most of proteins have evolved independently, and few of proteins are conserved among the four systems. I tried to superimpose phylogenetic trees of homologous proteins in the four energy metabolism systems to predict the phylogeny of the four energy metabolisms, and the phylogeny of aerobic respiration, denitrification and sulfur respiration was predicted.

The three level predictions were also applied to chemical neurotransmission system (Fig. 1.2). Neurotransmitters except for neuropeptides are produced by synthases in presynaptic cells, and the neurotransmitters are released to synapse. The neurotransmitters in the synapse are captured by receptors in postsynaptic cells, and the postsynaptic cells are activated. The neurotransmitters in the synapse are uptaken by transporters in the presynaptic cells or degradated in the synapse, and the chemical neurotransmission is inactivated. There are various kinds of neurotransmitters, and synthases, receptors and transporters for each neurotransmitter exist in chemical neurotransmission system. Therefore, each functional system for chemical neurotransmission can be defined as a system composed of synthases, receptors and transporters for each neurotransmission. I tried to predict how the chemical neurotransmission systems have been diversified by means of superimposing the phylogenetic trees of synthases, receptors and transporters. The phylogenetic trees of some synthases and some receptors are possible to superimpose, and the

phylogenetic trees of some receptors and some transporters are also possible to superimpose. These proteins might evolve together. But most of other proteins seem to have evolved independently from this study. Therefore, the unit of evolution is not system in chemical neurotransmission systems. I also did phylogenetic analyses of receptors at the domain level and the amino acid level. I inferred domain composition changes which had changed protein functions from the domain level analyses. Some domain changes seem to be essential for generating some receptors. The essential amino acid substitutions for changing ligand specificities were also predicted from the amino acid level analysis.

Based on the phylogenetic analyses of energy metabolism and chemical neurotransmission system, I carried out phylogenetic analysis of voltage-gated potassium channels at both; domain level and amino acid level (chapter 4). Voltage-gated ion channels are important for generating action potentials in postsynaptic neurons after chemical neurotransmission. Voltage-gated ion channels have the three states; resting, activation and inactivation (Fig. 1.3). I focused on inactivation, one of the major electrophysiological features, and predicted how the diversification of the inactivation had occurred. There are two kinds of inactivation in voltage-gated potassium channels. One is N-type inactivation which is sudden inactivation immediately after activation, the other one is C-type inactivation which is slow inactivation. Reffering to phylogenetic trees and domain compositions of voltage-gated potassium channels, domain composition changes seem not to affect inactivation differences. Previous studies (Murrell-Lagnado and Aldrich 1993a; Murrell-Lagnado and Aldrich 1993b) suggest that the specific chemical features of 20 N-terminal amino acids are important for generating N-type inactivation (Fig. 1.3). Therefore, I investigated the specific chemical features of 20 N-terminal amino acid sequences in voltage-gated potassium channels. The specific chemical features of the 20 amino acids which induce N-type inactivation are found in

the three subtypes out of the 21 subtypes. A small number of amino acid substitutions might produce the three N-type inactivation subtypes. The specific chemical features are not found in 20 N-terminal amino acids in the other five subclasses which can generate N-type inactivation. The amino acid substitutions which produce the five subtypes may different from those of the three subtypes which have the specific chemical features in the 20 N-terminal amino acids.

I did phylogenetic analyses at system and domain levels in electron transfer energy metabolism systems, did analyses at system, domain and amino acid levels in chemical neurotransmission systems, and did analyses at domain and amino acid levels in voltage-gated potassium channel family. From system level analysis, most of proteins in functional systems seem to be evolved independently. Although functions have been conserved in energy metabolism systems and chemical neurotransmission system, most of proteins in these systems have not been conserved. Tree superimpositions predict the phylogeny among the three systems in energy metabolism system and co-evolutions of some proteins in chemical neurotransmission system. From domain level analyses, I predicted how electron transferring mechanisms have changed through evolution in energy metabolism and how protein functions have changed through evolution in chemical neurotransmission system. In neurotransmission system, some domain compositions seem to be essential for generating some receptors. From amino acid level analyses, I predicted the essential amino acid substitutions for changing ligand binding specificities in chemical neurotransmission system and those for changing inactivation differences in voltage-gated potassium ion channel family. Domain composition changes seem to be slower than a small number of amino acid substitutions referring to phylogenetic analyses in this study. The combinations of the slower domain composition changes and the faster amino acid substitutions may have changed

protein functions and these combinations might produce divergence of protein functions. Functional systems may have been conserved their functions by the combinations and been diversified also by the combinations.

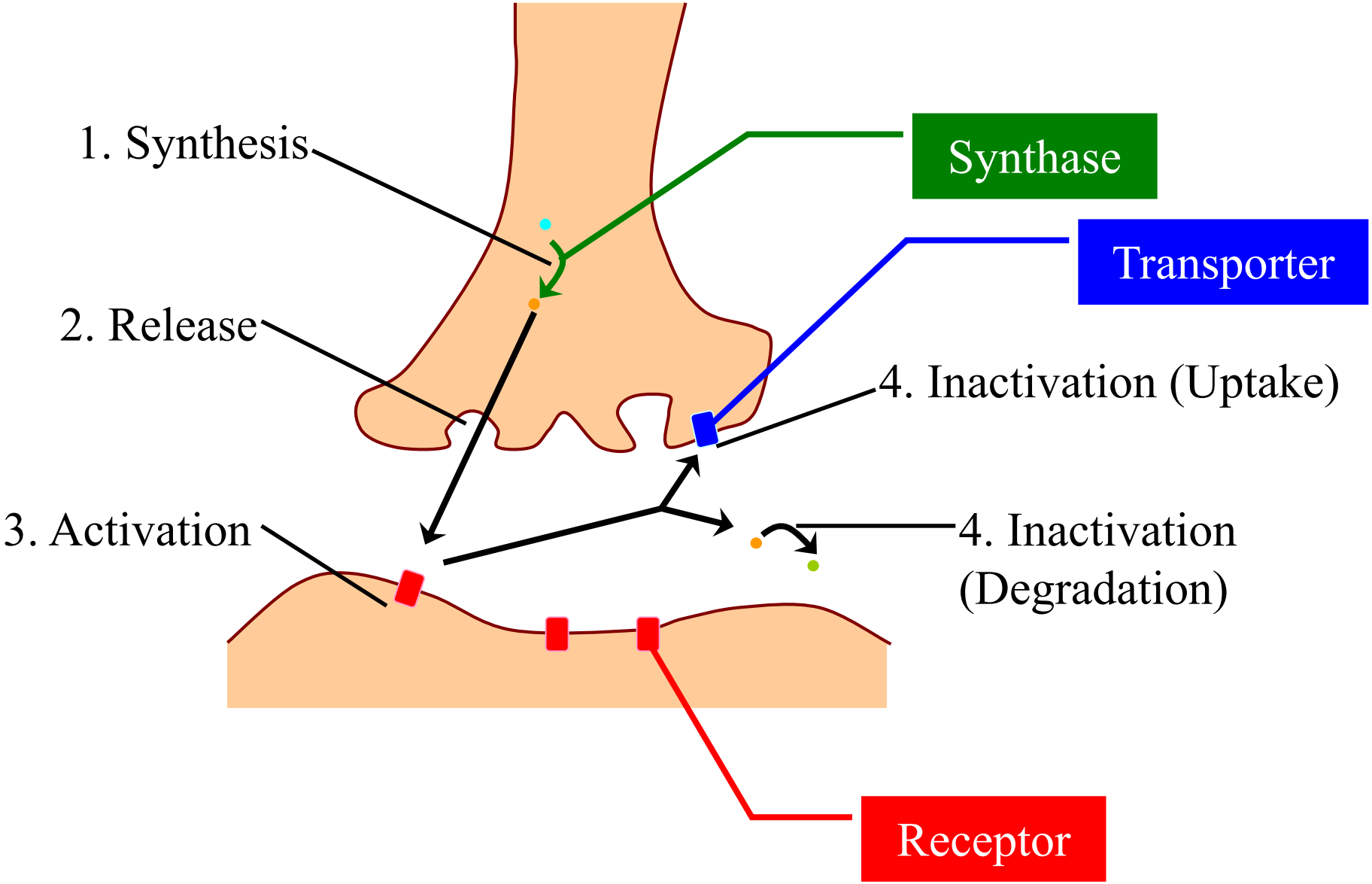


Fig. 1.2: Chemical neurotransmission. Neurotransmissions are synthesized by synthases (1), and stored into vesicles. When presynaptic neuron is activated, neurotransmitters are released to synapse (2). Neurotransmitters in synapse are captured by receptors, and postsynaptic neuron is activated (3). Neurotransmitters in synapse are uptaken by transporters or degraded by enzymes, and chemical neurotransmission is inactivated (4).

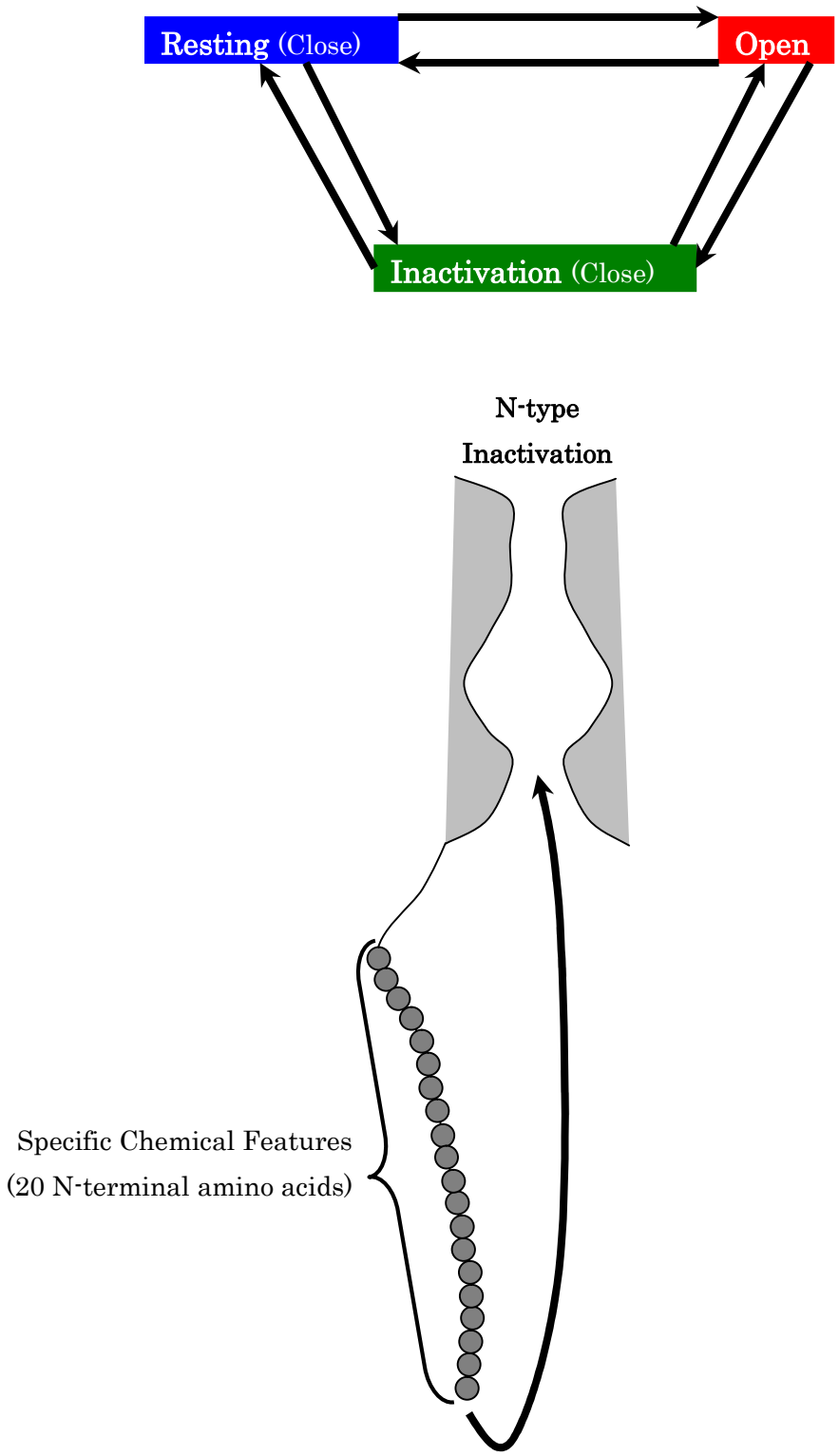


Fig. 1.3: Three states of ion channels. Pores of ion channel are shown. 20 N-terminal amino acids close a pore of voltage-gates ion channel.

2. Phylogenetic Analysis of Proteins in Energy Metabolism Systems

2.1 Introduction

Photosynthesis, aerobic respiration, denitrification, and sulfur respiration are energy generation metabolisms. Proton potential is generated through the electron transfer, and ATP synthase produces ATP with the help of this potential in all the four systems. Furthermore, the electron transporter in all the four metabolism systems is quinine, and cytochrome c works as an electron transporter in photosynthesis, aerobic respiration and denitrification. Some homologous proteins are used in different metabolic systems (Fig. 2.1). Cytochrome bc_1 complex in aerobic respiration and denitrification, and cytochrome b_6f complex in photosynthesis are evolutionarily related (Berry et al. 2000; Hauska et al. 1988). Subunit I and II of cytochrome c oxidase in aerobic respiration were shown to be homologous with nitric oxide reductase subunit b and monomeric nitrous oxide reductase in denitrification (Mogi et al. 1998; Saraste and Castresana 1994; Zumft et al. 1992). Between denitrification and sulfur respiration, respiratory nitrate reductase, periplasmic nitrogen reductase, and polysulfide reductase are homologous (Dias et al. 1999; Krafft et al. 1992; Moreno-Vivian et al. 1999).

It is of my interest how these proteins and characteristics common to the four energy metabolism systems evolved. Evolutionary relationships of some proteins in the four metabolisms were already analyzed as I described above. However, the exhaustive phylogenetic analysis of proteins among the four metabolisms has not been tried yet. I therefore conducted a molecular phylogenetic analysis of all available amino acid sequences of those four energy metabolism systems by using the homology search program PSI-BLAST (Altschul et al. 1997), because it can detect evolutionary relationships of highly diverged

proteins (Park et al. 1998). It is possible that homologous proteins in the four energy metabolism systems were also used in other nitrogen and sulfur metabolisms (nitrate assimilation, nitrogen fixation and nitrification in nitrogen metabolism, and assimilation and dissimilation in sulfur metabolism) in addition to denitrification and sulfur respiration. Therefore, I also included amino acid sequences in these metabolisms for my analysis.

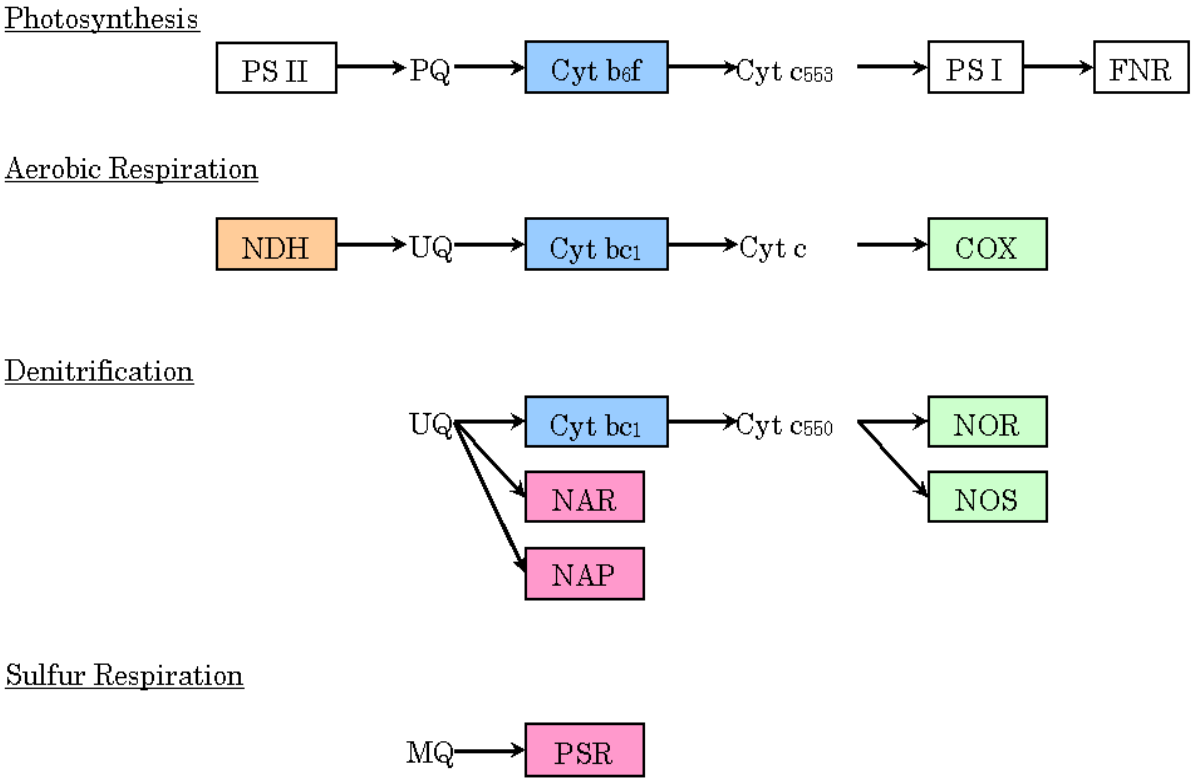


Fig. 2.1: Homologous proteins in photosynthesis, aerobic respiration, denitrification and sulfur respiration. Black arrows designate electron flows. Square objects are proteins. Same colored proteins are homologous. Abbreviations in this figure are as follows. PS II = photosystem II; PQ = plastoquinone; Cyt b_6f = cytochrome b_6f complex; Cyt c_{553} = cytochrome c_{553}; PS I = photosystem I; FNR = ferredoxin NADP reductase; NDH = NADH dehydrogenase; UQ = ubiquinone; Cyt bc_1 = cytochrome bc_1 complex; Cyt c = cytochrome c; COX = cytochrome c oxidase; NAR = respiratory nitrate reductase; NAP = periplasmic nitrate reductase; Cyt c_{550} = cytochrome c_{550}; NOR = nitric oxide reductase; NOS = nitrous oxide reductase; MQ = menaquinone; PSR = polysulfide reductase

2.2 Materials and Methods

2.2.1 Amino Acid Sequence Retrieval

I retrieved amino acid sequences of proteins involved in photosynthesis, aerobic respiration, nitrogen metabolism including denitrification, nitrate assimilation, nitrogen fixation and nitrification and sulfur metabolism including assimilation and dissimilation from SWISS-PROT (Boeckmann et al. 2003), by referring mainly to the pathway maps of oxidative phosphorylation, ATP synthesis, photosynthesis, nitrogen metabolism and sulfur metabolism in KEGG (http://www.genome.ad.jp/kegg/; Kanehisa et al. 2002). Fig. 2.2 shows the proteins used in this study (red-colored) involved in the four energy metabolism systems. Proteins whose amino acid sequence data are not available from SWISS-PROT are colored black in Fig. 2.2. The V-type ATP synthase in eukaryotic cells is included in the map of ATP synthesis in KEGG. However, it is not involved in this metabolism (Wieczorek et al. 1999). Therefore, this enzyme was excluded from the present study. In total, I used 182 proteins retrieved by this procedure (Table 2.1).

I used those sequence data as queries, and searched SWISS-PROT by using PSI-BLAST (Altschul et al. 1997) with SEG filter and 10 passes. Based on PSI-BLAST search result, I defined homologous groups of query sequences by taking their union. By this way, I can find amino acid regions (domains) shared by all the query sequences. Amino acid sequences retrieved through PSI-BLAST using these query sequences belonging to one homologous group are used for phylogenetic analysis. When function of a protein was not clearly defined, and was listed only as “hypothetical protein” in SWISSPROT, these sequences were ignored in the further analyses.

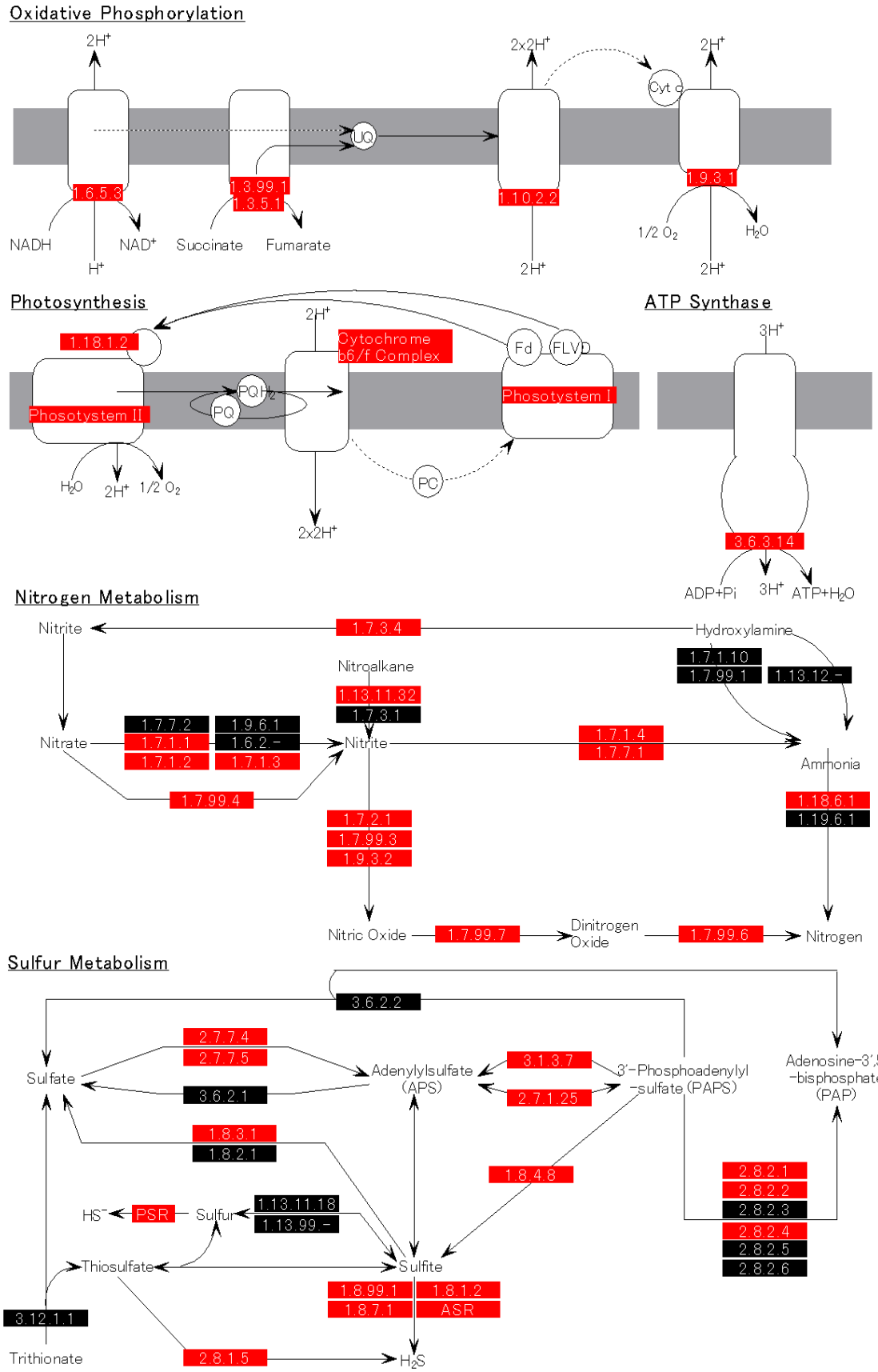


Fig. 2.2: Proteins I retrieved sequences (red colored proteins in this figure) mainly referring to pathway maps in KEGG.

Table 2.1: Amino acid sequences I used.

Metabolism	SWISS-PROT ID (Accession Number)
Photosynthesis	PSAA_SYNY3 (P29254)
	PSAB_SYNY3 (P29255)
	PSAC_SYNY3 (P32422)
	PSAD_SYNY3 (P19569)
	PSAE_SYNY3 (P12975)
	PSAF_SYNY3 (P29256)
	PSH1_ARATH (Q9SUI7)
	PSAI_SYNY3 (Q55330)
	PSAJ_SYNY3 (Q55329)
	PAK1_SYNY3 (P72712)
	PSAL_SYNY3 (P37277)
	PSAM_SYNY3 (P72986)
	PSAN_ARATH (P49107)
	PSAX_ANASP (P58566)
	PSB1_SYNY3 (P07826)
	PSBD_SYNY3 (P09192)
	PSBC_SYNY3 (P09193)
	PSBB_SYNY3 (P05429)
	PSBE_SYNY3 (P09190)
	PSBF_SYNY3 (P09191)
	PSBL_SYNY3 (Q55354)
	PSBJ_SYNY3 (P73070)
	PSBK_SYNY3 (P15819)
	PSBM_SYNY3 (P72701)
	PSBN_SYNY3 (P26286)
	PSBH_SYNY3 (P14835)
	PSBT_ARATH (P37259)
	PSBI_SYNY3 (Q54697)
	PSBO_SYNY3 (P10549)
	PSP1_ARATH (Q42029)
	PSBU_SYNY3 (Q55332)
	C55O_SYNY3 (Q55013)
	PSBW_SYNY3 (Q55356)
	PSBY_SYNY3 (P73676)
	PS11_SYNY3 (P74367)
	CYB6_SYNY3 (Q57038)
	PETD_SYNY3 (P27589)
	CYF_SYNY3 (P26287)
	UCRI_SYNY3 (P26290)
	PETL_ORYSA (P12180)
	PETM_SYNY3 (P74810)
	PETG_SYNY3 (P74149)
	PLAS_SYNY3 (P21697)
	FENR_SYNY3 (Q55318)
Photosynthesis, Aerobic Respiration	NUOA_ECOLI (P33597)
	NUOB_ECOLI (P33598)
	NUGC_SYNY3 (P19125)
	NUCC_SYNY3 (P27724)
	NUOE_ECOLI (P33601)
	NUOF_ECOLI (P31979)
	NUOG_ECOLI (P33602)
	NUOH_ECOLI (P33603)
	NUOI_ECOLI (P33604)

	NUOJ_ECOLI (P33605) NUOK_ECOLI (P33606) NUOL_ECOLI (P33607) NUOM_ECOLI (P31978) NUON_ECOLI (P33608)
Aerobic Respiration	DHSC_ECOLI (P10446) DHSD_ECOLI (P10445) DHSA_ECOLI (P10444) DHSB_ECOLI (P07014) FRDC_ECOLI (P03805) FRDD_ECOLI (P03806) COXX_YEAST (P21592) CYOD_ECOLI (P18403) COX3_YEAST (P00420) COX1_YEAST (P00401) COX2_YEAST (P00410) CX5A_YEAST (P00424) COX6_YEAST (P00427) COX4_YEAST (P04037) COXE_YEAST (P32799) COXG_YEAST (Q01519) COXH_HUMAN (P09669) COX7_YEAST (P10174) COXM_HUMAN (P24311) COX8_YEAST (P04039) COXR_HUMAN (P10176) COXZ_YEAST (P19516) COXW_YEAST (P40086) COXS_YEAST (Q12287)
Aerobic Respiration, Nitrogen Metabolism	UCRI_YEAST (P08067) CYB_YEAST (P00163) CY1_YEAST (P07143) UCR1_YEAST (P07256) UCR2_YEAST (P07257) UCRH_YEAST (P00127) UCR7_YEAST (P00128) UCRQ_YEAST (P08525) UCR9_YEAST (P22289) UCRX_YEAST (P37299)
Nitrogen Metabolism	HAO_NITEU (Q50925) 2NPD_NEUCR (Q01284) NARB_SYNY3 (P73448) NASA_KLEPN (Q06457) NARG_ECOLI (P09152) NARH_ECOLI (P11349) NARI_ECOLI (P11350) NAPA_ECOLI (P33937) NAPB_ECOLI (P33933) NIA1_ARATH (P11832) NIA_BETVE (P27783) NIA_NEUCR (P08619) NIRB_ECOLI (P08201) NIRD_ECOLI (P23675) NIR_SYNP7 (P39661) NIR_RHOSH (Q53239) NIRS_PARDE (Q51700) NORB_PSEST (P98008) NORC_PSEST (Q52527)

	NOSZ_PSEST (P19573)
	NIFD_AZOVI (P07328)
	NIFK_AZOVI (P07329)
	NIH1_AZOVI (P00459)
	VNFD_AZOVI (P16855)
	VNFK_AZOVI (P16856)
	VNFG_AZOVI (P16857)
	NIH2_AZOVI (P15335)
	ANFD_AZOVI (P16266)
	ANFK_AZOVI (P16267)
	ANFG_AZOVI (P16268)
	NIH3_AZOVI (P16269)
Sulfur Metabolism	CYSN_ECOLI (P23845)
	CYSD_ECOLI (P21156)
	SAT_SYNY3 (P74241)
	APA1_YEAST (P16550)
	DPNP_ARATH (Q42546)
	CYSC_ECOLI (P23846)
	SUP1_HUMAN (P50225)
	SUHA_HUMAN (Q06520)
	SUOE_HUMAN (P49888)
	CYSH_ECOLI (P17854)
	DSRA_ARCFU (Q59109)
	DSRB_ARCFU (Q59110)
	CYSI_ECOLI (P17846)
	CYSJ_ECOLI (P38038)
	SIR_SYNY3 (P72854)
	ASRA_SALTY (P26474)
	ASRB_SALTY (P26475)
	ASRC_SALTY (P26476)
	PHSA_SALTY (P37600)
	PHSB_SALTY (P37601)
	PHSC_SALTY (P37602)
	PSRA_WOLSU (P31075)
	PSRB_WOLSU (P31076)
	PSRC_WOLSU (P31077)
ATP Synthase	ATPE_ECOLI (P00832)
	ATPB_ECOLI (P00824)
	ATPG_ECOLI (P00837)
	ATPA_ECOLI (P00822)
	ATPD_ECOLI (P00831)
	ATPF_ECOLI (P00859)
	ATPL_ECOLI (P00844)
	ATP6_ECOLI (P00855)
	ATPE_YEAST (P21306)
	ATPD_YEAST (Q12165)
	ATPB_YEAST (P00830)
	ATPG_YEAST (P38077)
	ATPA_YEAST (P07251)
	ATPO_YEAST (P09457)
	ATP9_YEAST (P00841)
	ATP6_YEAST (P00854)
	ATPF_YEAST (P05626)
	ATPJ_HUMAN (P56385)
	ATPR_HUMAN (P18859)
	ATPK_HUMAN (P56134)
	ATP8_YEAST (P00856)
	ATP7_YEAST (P30902)

ATPK_YEAST	(Q06405)
AT14_YEAST	(Q12349)
AT18_YEAST	(P81450)
AT19_YEAST	(P81451)
ATPN_YEAST	(Q12233)
VATC_METJA	(Q57672)
VATF_METJA	(Q57671)
VATA_METJA	(Q57670)
VATB_METJA	(Q57669)
VATE_METJA	(Q57673)
ATPL_METJA	(Q57674)
VATI_METJA	(Q57675)
VATD_METJA	(Q58032)

2.2.2 Phylogenetic Analysis

Because amino acid sequences which I analyzed are so diverged, I did not use the whole region of each protein for phylogenetic analysis so as to obtain reliable multiple alignments and phylogenetic trees. I defined conserved regions based on domains listed in Pfam database (http:// www.sanger.ac.uk/Software/Pfam/; Bateman et al. 2002). Pfam is a database of conserved amino acid sequence regions. Each conserved region is called a family. Multiple alignment of conserved region is constructed in each family, and is available from the Pfam web site. I used multiple alignments available in the Pfam database. When I encountered amino acid regions which are not defined by Pfam but were weakly homologous with a known Pfam domain, those amino acid sequences and corresponding Pfam domain sequences were multiply aligned by using ClustalX (Thompson et al. 1997). Neighbor-joining trees (Saitou and Nei 1987) were constructed by using MEGA2 (Kumar et al. 2001). Numbers of amino acid substitutions were computed both for Poisson correction and gamma distance with shape parameter 1 (default value).

2.2.3 Identification of Electron Transfer Patterns

In this study, I analyzed the relationship between amino acid sequences and catalytic features mainly focused on patterns of electron transfer, because electron transfer is essential for catalysis. Information on catalytic features especially about patterns of electron transfer were retrieved from both SWISS-PROT database and bibliographies cited in this paper. I categorized proteins whose amino acid sequence data were used in this study into 4 categories as follows. Category of each amino acid sequence data is shown in Table 2.2.

A. The amino acid sequence of the protein was experimentally determined, and its function is known. Moreover, electron transfer pattern in the protein has been

revealed or predicted. References regarding its function are shown in Table 2.2.

B. The amino acid sequence of the protein was experimentally determined, and its function is known. References regarding its function are shown in Table 2.2.

C. The amino acid sequence of the protein was predicted only by translating genomic or cDNA nucleotide sequence, and its function was predicted by sequence similarity in SWISS-PROT.

D. The amino acid sequence of the protein was predicted only by translating genomic or cDNA nucleotide sequence, and there is no description on its function in SWISS-PROT.

Table 2.2: Electron transfer patterns of proteins found to belong to three homologous groups.

Group 1

Enzyme	SWISS-PROT ID	Accession Number	Cate-gory	Reference
NADH Dehydrogenase	NUOG_ECOLI	P33602	C	
	NUOG_RICPR	Q9ZCF6	C	
	NUOG_SALTY	P33900	C	
	NUOG_STRCO	Q9XAR0	C	
	NUG2_RHIME	P56914	C	
	NUAM_ACACA	Q37373	C	
	NUAM_BOVIN	P15690	A	Ohnishi 1998
	NUAM_HUMAN	P28331	B	Chow et al. 1991
	NUAM_RECAM	O21241	C	
	NUAM_SOLTU	Q43644	B	Rasmusson et al. 1998
	NUAM_NEUCR	P24918	B	Preis et al. 1991
	NQO3_PARDE	P29915	B	Xu et al. 1992
	NQO3_THETH	Q56223	C	
Formate Dehydrogenase	FDHA_METFO	P06131	B	Shuber et al. 1986
	FDHA_METJA	Q60314	C	
	FDHF_ECOLI	P07658	A	Boyington et al. 1997; Sauter et al. 1992
	FDOG_ECOLI	P32176	B	Abaibou et al. 1995
	FDNG_ECOLI	P24183	A	Jormakka et al. 2002
	FDXG_HAEIN	P46448	C	
Nitrate Reductase	NAPA_DESDE	P81186	A	Dias et al. 1999
	NAPA_ALCEU	P39185	B	Siddiqui et al. 1993
	NAPA_ECOLI	P33937	B	Grove et al. 1996; Thomas et al. 1999
	NAPA_PARPN	Q56350	A	Berks et al. 1995; Breton et al. 1994; Butler et al. 1999; Moreno-Vivian et al. 1999
	NAPA_RHOSH	Q53176	B	Reyes et al. 1996
	NARB_SYNP7	P39458	B	Rubio et al. 1996
	NARB_SYNY3	P73448	B	Aichi et al. 2001
	NASA_KLEOX	Q06457	B	Lin et al. 1993

	NASC_BACSU	P42434	B	Ogawa et al. 1995
	NARG_BACSU	P42175	C	
	NARG_ECOLI	P09152	A	Blasco et al. 1989; Guigliarelli et al. 1992; Magalon et al. 1997; Magalon et al. 1998; Moreno-Vivian et al. 1999
	NARZ_ECOLI	P19319	B	Blasco et al. 1990
Polysulfide Reductase	PSRA_WOLSU	P31075	B	Krafft et al. 1992
Thiosulfate Reductase	PHSA_SALTY	P37600	B	Heinzinger et al. 1995
Dimethyl Sulfoxide Reductase	DMSA_ECOLI	P18775	A	Cammack and Weiner 1990; Heffron et al. 2001; Rothery et al. 1999; Rothery and Weiner 1996; Trieber et al. 1996
	DMSA_HAEIN	P45004	C	
	DMSA_RHOCA	Q52675	A	Bray et al. 2000; McAlpine et al. 1998; Schneider et al. 1996; Stewart et al. 2000
	DMSA_RHOSH	Q57366	A	Li et al. 2000; Schindelin et al. 1996
Trimethylamine-N-oxide Reductase	TORA_ECOLI	P33225	A	Gon et al. 2001
	TORA_SHEMA	O87948	B	Dos Santos et al. 1998
	TORZ_ECOLI	P46923	B	Gon et al. 2000
	TORZ_HAEIN	P44798	C	
Biotin Sulfoxide Reductase	BISC_ECOLI	P20099	B	Pierson and Campbell 1990
	BISC_RHOSH	P54934	A	Garton et al. 2000; Pollock and Barber 2001; Temple et al. 2000

Group 2

Enzyme	SWISS-PROT ID	Accession Number	Cate-gory	Reference
Sulfite Reductase [NADPH] Hemoprotein	CYSI_ECOLI	P17846	A	Crane et al. 1995; Crane et al. 1997b; Crane et al. 1997c
	CYSI_SALTY	P17845	B	Ostrowski et al. 1989b
	CYSI_THIRO	P52673	B	Bruhl et al. 1996
Sulfite Reductase (Ferredoxin)	SIR_SYNY3	P72854	C	
	SIR_SYNP7	P30008	B	Gisselmann et al. 1993
	-	Q42590	B	Bruhl et al. 1996
Sulfite Reductase, Assimilatory-type	SIR_DESVH	Q05805	A	Tan and Cowan 1991
Ferredoxin--nitrite Reductase	NIR_SYNP7	P39661	C	
	NIR_PHOLA	Q51879	B	Merchan et al. 1995
	NIR_SPIOL	P05314	B	Back et al. 1988; Back et al. 1991
	NIR_BETVE	P38500	C	
Nitrite Reductase [NAD(P)H]	NIR_NEUCR	P38681	A	Lafferty and Garrett 1974; Prodouz and Garrett 1981; Vega 1976
	NIR_EMENI	P22944	B	Johnstone et al. 1990
	NIRB_KLEOX	Q06458	B	Lin et al. 1993
	NASD_BACSU	P42435	B	Ogawa et al. 1995
	NIRB_ECOLI	P08201	B	Jayaraman et al. 1987; Peakman et al. 1990
Anaerobic Sulfite Reductase	ASRC_SALTY	P26476	B	Huang and Barrett et al. 1991
Sulfite Reductase, Dissimilatory-type	DSRA_ARCFU	Q59109	B	Dahl et al. 1993
	DSRA_CHRVI	O33998	B	Hipp et al. 1997
	DSRB_ARCFU	Q59110	B	Dahl et al. 1993
	DSVA_DESVH	P45574	B	Karkhoff-Schweizer et al. 1995; Pierik et al. 1992
	DSVB_DESVH	P45575	B	Karkhoff-Schweizer et al. 1995; Pierik et al. 1992

Group 3

Enzyme	SWISS-PROT ID	Accession Number	Cate-gory	Reference
CDP-6-deoxy-delta-3,4-glucoseen Reductase	ASCD_YERPE	P37911	A	Gassner et al. 1996; Johnson et al. 1996
Anaerobic Sulfite Reductase	ASRB_SALTY	P26475	B	Huang and Barrett et al. 1991
Benzoate 1,2-dioxygenase	BENC_ACICA	P07771	A	Karlsson et al. 2002
Bifunctional P-450:NADPH-P450 Reductase	CPXB_BACME	P14779	A	Haines et al. 2001; Hazzard et al. 1997; Li and Poulos 1997; Peterson et al. 1997; Ravichandran et al. 1993; Sevrioukova et al. 1996; Sevrioukova et al. 1999a; Sevrioukova et al. 1999b; Yeom et al. 1995
Sulfite Reductase [NADPH] Flavoprotein	CYSJ_ECOLI	P38038	A	Eschenbrenner et al. 1995a; Eschenbrenner et al. 1995b; Gruez et al. 2000; Ostrowski et al. 1989a; Siegel et al. 1973
	MT10_YEAST	P39692	B	Hansen et al. 1994
Phenol Hydroxylase P5 Protein	DMPP_PSESP	P19734	A	Powlowski and Shingler 1990
Flavohemoprotein	FHP_CANNO	Q03331	B	Iwaasa et al. 1992
	FHP_YEAST	P39676	B	Zhu and Riggs 1992
	HMPA_ALCEU	P39662	B	Cramm et al. 1994
NADH Oxidoreductase hcr	HCR_ECOLI	P75824	A	van den Berg et al. 2000
Fruit Protein PKIWI502	K502_ACTCH	P43394	D	
Flavin Reductase	LUXG_VIBHA	P16447	C	
NADH-cytochrome b5 Reductase	MCR1_YEAST	P36060	B	Hahne et al. 1994
	NC5R_BOVIN	P07514	B	Ozols et al. 1984; Ozols et al. 1985; Strittmatter et al. 1992
	NC5R_RAT	P20070	B	Pietrini et al. 1988; Zenno et al. 1990
	NC5R_YEAST	P38626	C	
Methane Monooxygenase	MMOC_METCA	P22868	A	Lund et al. 1985
NADPH-cytochrome P450 Reductase	NCPR_HUMAN	P16435	B	Haniu et al. 1989
	NCPR_RAT	P00388	B	Porter and Kasper 1985; Porter et al. 1990
	NCPR_YEAST	P16603	B	Yabusaki et al. 1988
Naphthalene 1,2-dioxygenase System Ferredoxin-NAD(+) Reductase Component	NDOR_PSEPU	Q52126	B	Simon et al. 1993
Nitrate Reductase	NIA_ASPNG	P36858	B	Unkles et al. 1992
	NIA_PHYIN	P39864	B	Pieterse et al. 1995
	NIA_PICAN	P49050	B	Avila et al. 1995
	NIA2_ARATH	P11035	A	Skipper et al. 2001
Nitric-oxide Synthase	NOS_DROME	Q27571	B	Regulski and Tully 1995
	NOS1_RAT	P29476	B	Bredt et al. 1991
	NOS2_MOUSE	P29477	A	Aoyagi et al. 2001; Crane et al. 1997a; Crane et al. 1998; Crane et al. 1999; Ghosh et al. 1999; McMillan et al. 2000; Siddhanta et al. 1998
	NOS3_BOVIN	P29473	B	Lamas et al. 1992

	NOS3_HUMAN	P29474	B	Janssens et al. 1992
	NS2A_HUMAN	P35228	A	Fischmann et al. 1999; Li et al. 1999
Phthalate Dioxygenase Reductase	PDR_BURCE	P33164	A	Batie et al. 1991; Correll et al. 1992
	PHT2_PSEPU	Q05182	B	Nomura et al. 1992
Phenoxybenzoate Dioxygenase	POBB_PSEPS	Q52186	B	Dehmel et al. 1995
	RFBI_SALTY	P26395	D	
Toluene-4-Monooxygenase	TMOF_PSEME	Q03304	B	Yen and Karl 1992
Vanillate O-demethylase Oxidoreductase	VANB_PSEPU	O54037	B	Venturi et al. 1998
Xylene Monooxygenase	XYLA_PSEPU	P21394	A	Shaw and Harayama 1992
Toluate 1,2-dioxygenase	XYLZ_PSEPU	P23101	B	Harayama et al. 1991
Putative Dioxygenase	YEAX_ECOLI	P76254	C	
Ferredoxin-NADP Reductase	FENR_ANASO	P21890	A	Arakaki et al. 1997; Martinez-Julvez et al. 1998; Martinez-Julvez et al. 2001; Mayoral et al. 2000; Morales et al. 2000; Serre et al. 1996
	FENR_SYNY3	Q55318	A	Arakaki et al. 1997; van Thor et al. 1999
	FENR_PEA	P10933	B	Newman and Gray 1988
	FENR_SPIOL	P00455	B	Jansen et al. 1988; Karplus et al. 1984
	FENR_AZOVI	Q44532	A	Arakaki et al. 1997; Isas and Burgess 1994; Sridhar Prasad et al. 1998
	FENR_BUCAP	Q9Z615	C	
	FENR_ECOLI	P28861	A	Arakaki et al. 1997; Bianchi et al. 1993; Ingelman et al. 1997

2.3 Results

2.3.1 Homologous Groups

I found 24 homologous groups of proteins through the PSI-BLAST search. Because my objective is to discover the changes of metabolic pattern among the four energy metabolism systems and the additional nitrogen and sulfur metabolisms, I only used homologous groups including proteins belonging to more than one metabolism system in the further analyses. Homologous groups suitable for my objective were groups 1-8. Table 2.3 shows the list of these eight groups. Homologous relationships of groups 5-8 were already reported (Berry et al. 2000; Hauska et al. 1988; Mogi et al. 1998; Saraste and Castresana 1994). All the proteins of group 4 had a conserved domain defined as Fer4 domain (4Fe-4S ferredoxin binding domain) by Pfam database (Bateman et al. 2002). Unfortunately, this domain is very short (less than 50 amino acids), and I could not construct a reliable phylogenetic tree. I thus analyzed groups 1-3 in the followings.

Table 2.3: Eight homologous protein groups found in this study

Group 1

ID	Name	Metabolism	No. of homologous proteins
NUOG_ECOLI	NADH Dehydrogenase Chain G	O	44
NARG_ECOLI	Respiratory Nitrate Reductase Alpha Chain	N_{DN}	
NARB_SYNY3	Assimilatory Nitrate Reductase (Ferredoxin-dependent)		
NASA_KLEOX	Assimilatory Nitrate Reductase (NADH-dependent)		
NAPA_ECOLI	Periplasmic Nitrate Reductase		
PHSA_SALTY	Thiosulfate Reductase PhsA Subunit	S_{DS}	
PSRA_WOLSU	Polysulfide Reductase Chain A	S_R	

Group 2

ID	Name	Metabolism	No. of homologous proteins
NIR_SYNP7	Ferredoxin--nitrite Reductase	N_{AS}	21
NIRB_ECOLI	Nitrite Reductase (NAD(P)H) Large Subunit		
DSRA_ARCFU	Sulfite Reductase (Dissimilatory-type) Alpha Subunit	S_{DS}	
DSRB_ARCFU	Sulfite Reductase (Dissimilatory-type) Beta Subunit		
ASRC_SALTY	Anaerobic Sulfite Reductase Subunit C	S_{AS}	
CYSI_ECOLI	Sulfite Reductase (NADPH) Beta-component		
SIR_SYNY3	Sulfite Reductase (Ferredoxin)		

Group 3

ID	Name	Metabolism	No. of homologous proteins
FENR_SYNY3	Ferredoxin--NADP Reductase	P	122
NIA2_ARATH	Nitrate Reductase [NADH]	N_{AS}	
NIA_BETVE	Nitrate Reductase [NAD(P)H]		
NIA_NEUCR	Nitrate Reductase [NADPH]		
CYSJ_ECOLI	Sulfite Reductase (NADPH) Alpha-component	S_{AS}	
ASRB_SALTY	Anaerobic Sulfite Reductase Subunit B		

Group 4

ID	Name	Metabolism	No. of homologous proteins
PSAC_SYNY3	Photosystem I Subunit VII (PsaC)	P	185
NUOI_ECOLI	NADH Dehydrogenase Chain I	O	
NARH_ECOLI	Respiratory Nitrate Reductase Beta Chain	N_{DN}	
PSRB_WOLSU	Polysulfide Reductase Chain B	S_R	
PHSB_SALTY	Thiosulfate Reductase PhsB Subunit	S_{DS}	
DSRA_ARCFU	Sulfite Reductase (Dissimilatory-type) Alpha Subunit		
DSRB_ARCFU	Sulfite Reductase (Dissimilatory-type) Beta Subunit		
ASRC_SALTY	Anaerobic Sulfite Reductase Subunit C	S_{AS}	

Group 5

ID	Name	Metabolism	No. of homologous proteins
CYB6_SYNY3	Cytochrome b6f Complex Cytochrome b6 Subunit	P	956
CYB_YEAST	Cytochrome bc1 Complex Cytochrome b Subunit	O, N_{DN}	

Group 6

ID	Name	Metabolism	No. of homologous proteins
UCRI_SYNY3	Cytochrome b6f Complex Iron-sulfur Subunit	P	37
UCRI_YEAST	Cytochrome bc1Complex Iron-sulfur Subunit	O, N_{DN}	

Group 7

ID	Name	Metabolism	No. of homologous proteins
COX2_YEAST	Cytochrome c Oxidase Polypeptide II	O	247
NOSZ_PSEST	Nitrous-oxide Reductase	N_{DN}	

Group 8

ID	Name	Metabolism	No. of homologous proteins
PETD_SYNY3	Cytochrome b6f Complex Subunit 4	P	937
CYB_YEAST	Cytochrome bc1 Complex Cytochrome b Subunit	O, N_{DN}	

P = photosynthesis; O = aerobic respiration; N_{DN} = denitrification; N_{AS} = nitrate assimilation; S_R = sulfur respiration; S_{DS} = dissimilation in sulfur metabolism; S_{AS} = assimilation in sulfur metabolism

2.3.1.1 Group 1

Domain compositions of 44 proteins belonging to homologous group 1 are shown in Fig. 2.3. All the group 1 proteins have the Molybdopterin domain that has a binding site to molybdenum (Boyington et al. 1997, Czjzek et al. 1998, Dias et al. 1999, Jormakka et al. 2002, Li et al. 2000, McAlpine et al. 1998, Schindelin et al. 1996, Schneider et al. 1996, Stewart et al. 2000).

The phylogenetic tree of this group is presented in Fig. 2.4, based on the multiple alignment of the Molybdopterin domain. Although the 9 sequences (FDXG_HAEIN, FDOG_ECOLI, FDNG_ECOLI, NQO3_THETH, NUG2_RHIME, NUOG_MYCTU, NUOG_ECOLI, NUOG_SALTY and NUOG_STRCO) listed in Fig. 2.3 have Molybdopterin domain, those domain sequences were fragmentary. I therefore excluded the 9 sequences from the phylogenetic tree (Fig. 2.4). Topologies of two NJ trees constructed by using Poisson correction (not shown) and gamma distance (Fig. 2.4) were slightly different. I thus constructed a tree (not shown) by using the combination of Molybdopterin domain and Molydop_binding domain regions. Although 9 proteins lack Molydop_binding domain, they formed a clear monophyletic clusters with a high bootstrap probability (see Fig. 2.4), and exclusion of those sequences does not influence the overall tree topology. This new tree is more reliable than the tree constructed by using only Molybdopterin domain because of higher bootstrap probabilities (shown in parentheses in Fig. 2.4). The topology of this new tree was consistent with the tree computed for the gamma distance, thus I chose that tree as shown in Fig. 2.4.

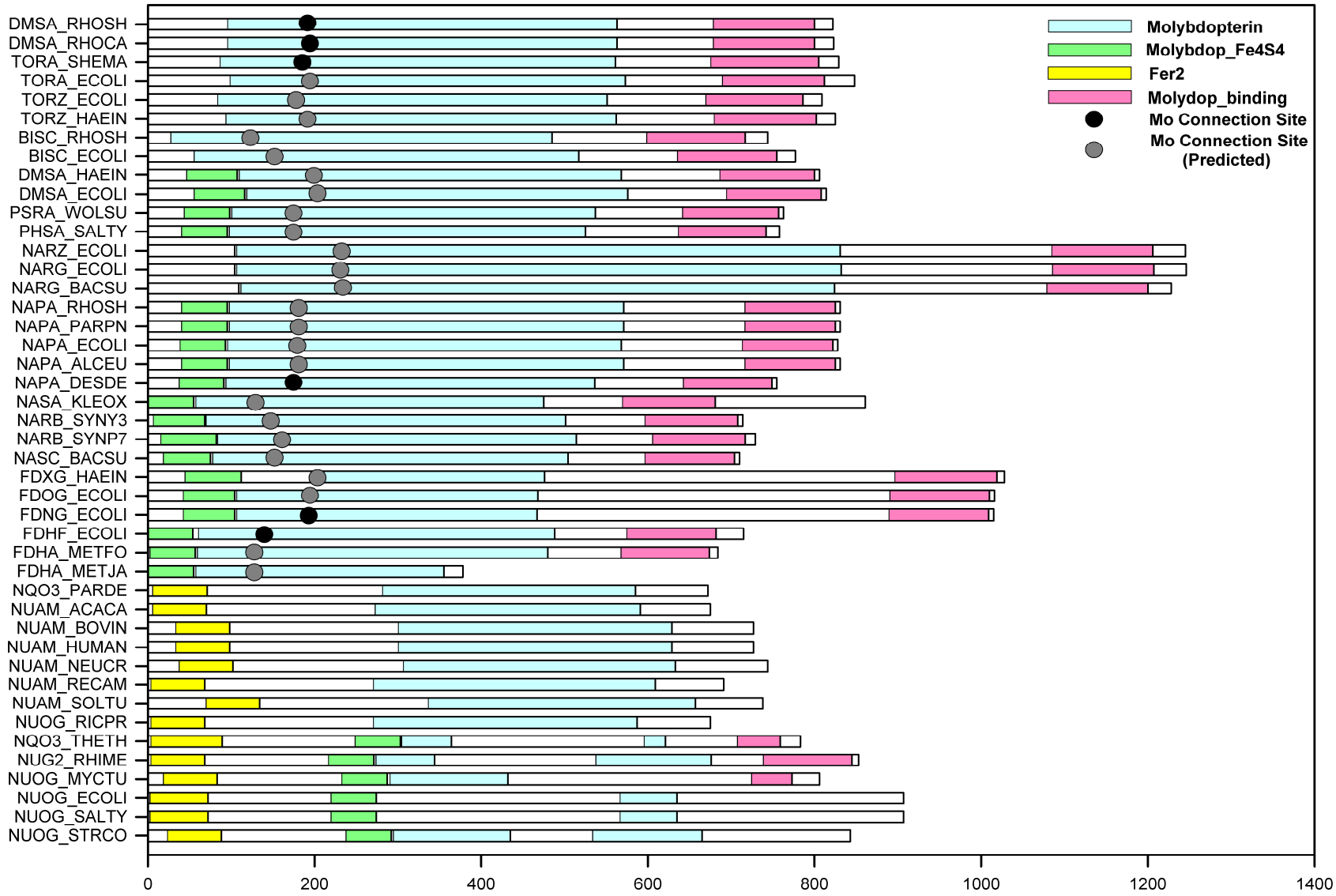


Fig. 2.3: Domain structures of group 1 proteins. SWISS-PROT IDs were used for protein names. Domains are defined following the Pfam database. Binding sites of molybdenum (Mo) which is important for catalysis are shown by full circles. I predicted molybdenum binding sites based on the multiple alignment of group 1 proteins. I regarded amino acid which is the same column of known molybdenum binding site as molybdenum binding site (shown by gray circles).

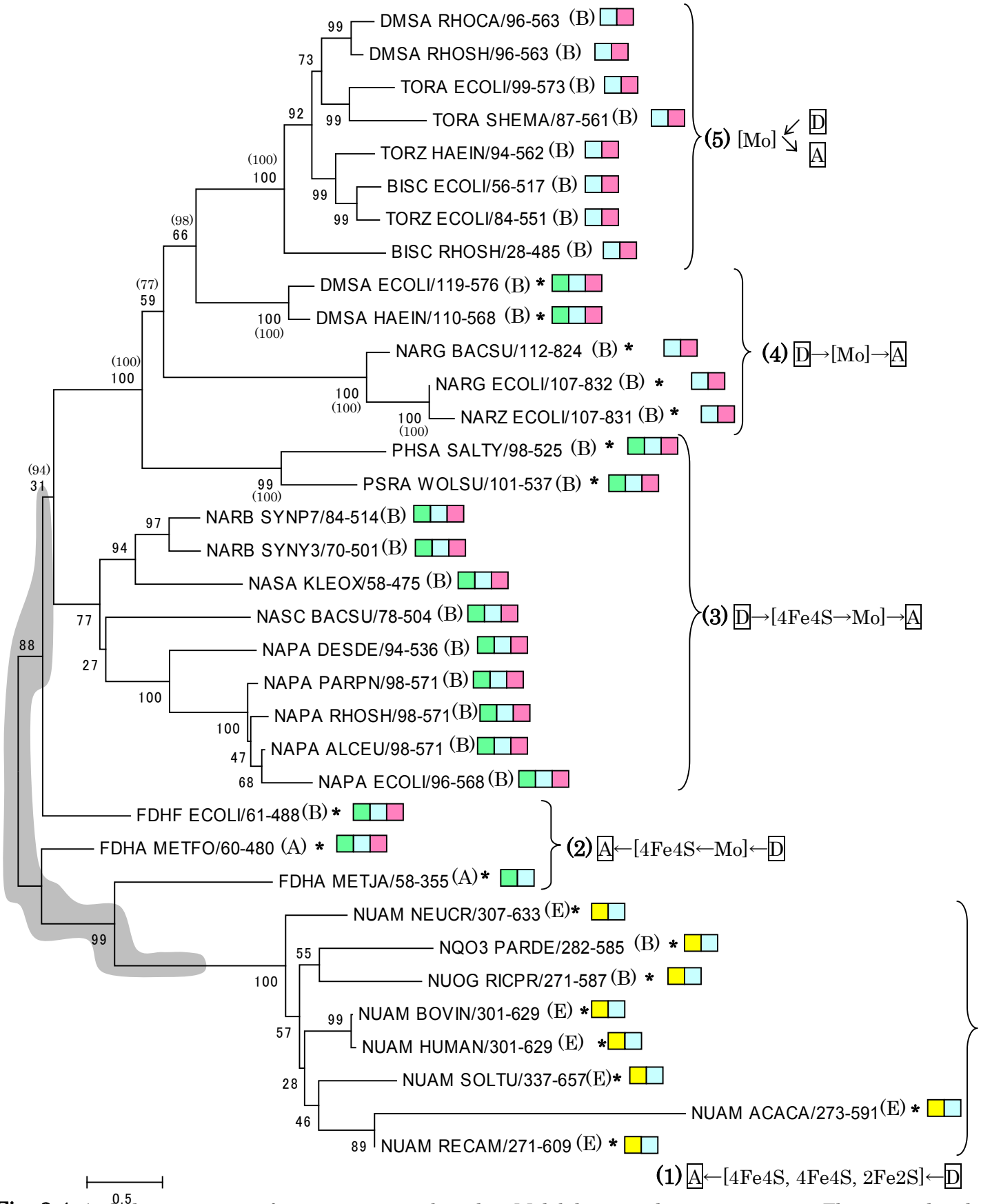


Fig. 2.4: A phylogenetic tree of group 1 proteins based on Molybdopterin domain sequences. The gray colored region shows possible locations of root for this tree. Each colored square designates a Pfam domain. Explanatory note about color is shown in Fig. 2.3. Electron transfer patterns are also shown. Arrows of the patterns indicate the direction of electron flow. D and A surrounded by square mean electron donor and acceptor, respectively. Unknown patterns of electron transfer were predicted from known patterns of evolutionarily closely related proteins. Proteins with * can generate proton gradient. Amino acid sequences of eubacteria, archaea and eukaryotes are designated by (B), (A) and (E), respectively.

All the group 1 proteins are involved in enzymatic reactions accompanied by electron transfer. Group 1 proteins shown in Fig. 2.4 can be divided into five types based on domain structures corresponding to patterns of electron transfer. Arrows in electron transfer patterns of Fig. 2.4 designate directions of electron flows. Electron flow in one pair of brackets occurs in each group 1 protein. Although electron transfer patterns of type 4 and 5 look same in Fig. 2.4, electron transfer patterns among electron donor, molybdenum and electron acceptor seem to be different between type 4 and 5 proteins. Type 4 protein gets electron from electron donor (substrate) at active site and seems to give it to electron acceptor (another subunit) by way of another side (Rothery et al. 1999). But type 5 gets electron from electron donor (substrate) at active site and probably gives it to at the same active site (Schindelin et al. 1996).

Type 1 proteins (NADH dehydrogenase subunit) do not bind molybdenum. Molybdenum plays an important role for reaction with substrates involving electron transfer. The main function of type 1 proteins is electron transfer from one subunit to another subunit through three ferredoxins (two 4Fe4Ss and one 2Fe2S) (Ohnishi 1998). Only type 1 protein has Fer2 domain, 2Fe2S ferredoxin binding domain, among group 1 proteins, and one ferredoxin is bound to this domain.

Type 2 proteins (formate dehydrogenase subunit) bind molybdenum. Electron flow of this protein is as follows; electron donor (substrate [formate]) => molybdenum => 4Fe4S => electron acceptor (another subunit) (Boyington et al. 1997; Jormakka et al. 2002). FDHA_METJA lost Molybdop_binding domain that is important for binding molybdenum (Boyington et al. 1997, Czjzek et al. 1998, Dias et al. 1999, Jormakka et al. 2002, Schindelin et al. 1996, Schneider et al. 1996). The function of FDHA_METJA was predicted only by sequence similarity (Table 2.3). FDHA_METJA may not bind molybdenum and lose the catalytic function by molybdenum. Type 3 proteins (subunit of assimilatory and periplasmic

nitrate reductase) have the same domains as type 2 proteins except for FDHA_METJA, but the electron flow is reversed (Breton et al. 1994; Butler et al. 1999; Dias et al. 1999). 4Fe4S ferredoxin is bound to Molybdop_Fe4S4 domain, 4Fe4S ferredoxin binding domain, in proteins of types 2 and 3.

Type 4 proteins (subunit of trimer type DMSO reductase and respiratory nitrate reductase) do not bind 4Fe4S ferredoxin. Molybdop_4Fe4S domain of trimer type DMSO reductase subunit (DMSA_ECOLI and DMSA_HAEIN) is less conserved, and that of respiratory nitrate reductase subunit (NARG_BACSU, NARG_ECOLI and NARZ_ECOLI) was lost. Four cysteines in this domain are essential for binding 4Fe4S ferredoxin. The arrangement of the four cysteines in type 4 is different from that of types 2 and 3 (Trieber et al. 1996). There are two amino acids between the first N-terminus two cysteines in type 2 and 3 proteins, but there are three amino acids in trimer type DMSO reductase subunit. In respiratory nitrate reductase subunit, the first cysteine is replaced by histidine. Changed cysteine arrangement in type 4 lost the ability for binding ferredoxin (Magalon et al. 1998; Trieber et al. 1996). The cysteine arrangement is not found in type 5 proteins (biotin sulfoxide reductase, trimethylamine-N-oxide reductase and monomer type DMSO reductase), therefore type 5 does not bind ferredoxin.

Which type of electron transfer pattern was ancestral among the five types in group 1 proteins? Type 2 proteins exist in eubacteria and archaea. It suggests that type 2 proteins emerged before the divergence of eubacteria and archaea. Types 3-5 proteins are found only in eubacteria (see Fig. 2.4), and it suggests their relatively recent origin, after the divergence of eubacteria and archaea. Type 1 proteins exist both in eubacteria and eukaryotes. Type 1 protein gene in some eukaryote species (NUAM_ACACA and NUAM_RECAM) is encoded in mitochondrial genome. It is hypothesized that mitochondria was established as a result of

endosymbiosis of an ancestor of alpha proteobacteria *P. denitrificans*, and is also hypothesized that type 1 protein gene encoded in mitochondrial genome was transferred to eukaryotic genome after the endosymbiosis (Finel. 1998). Andersson et al. (1998) showed that alpha proteobacteria *R. prowazekii* is the closest eubacterial relative of mitochondria. Therefore, type 1 protein gene of eukaryotes could be derived from that of the ancestor of alpha proteobacteria. Consequently, type 2 proteins seem to be the ancestor type protein in group 1. The gray colored region in Fig. 2.4 shows possible locations of root for this tree. I couldn't decide the exact root point.

2.3.1.2 Group 2

The domain structures of 21 proteins belonging to group 2 are shown in Fig. 2.5. All the group 2 proteins have NIR_SIR domain and NIR_SIR_ferr domain defined by Pfam in adjacent locations. Both domains have siroheme binding sites, and confusingly NIR_SIR domain have ferredoxin binding sites according to Pfam annotation (Crane et al. 1995). Hatched domains in Fig. 2.5 are not defined by Pfam, but they were detected as weakly similar to these two Pfam domains through PSI-BLAST search. Some of them were previously noticed by Crane and Getzoff (1996) and Larsen et al (1999). I constructed a phylogenetic tree of group 2 proteins by using the conserved regions consisting of the two domains (Fig. 2.6). The first 9 proteins of Fig. 2.5 have two NIR_SIR_ferr and two NIR_SIR domains, although ferredoxin binding sites are lost in hatched NIR_SIR domains (Crane et al. 1995). These two NIR_SIR_ferr and two NIR_SIR domains suggest internal gene duplication. In the case of these duplicated conserved regions, I used each region as independent sequence. For example, I used two sequences in different conserved regions of CYSI_ECOLI; amino acid positions from 109 to 233 and those from 387 to 489.

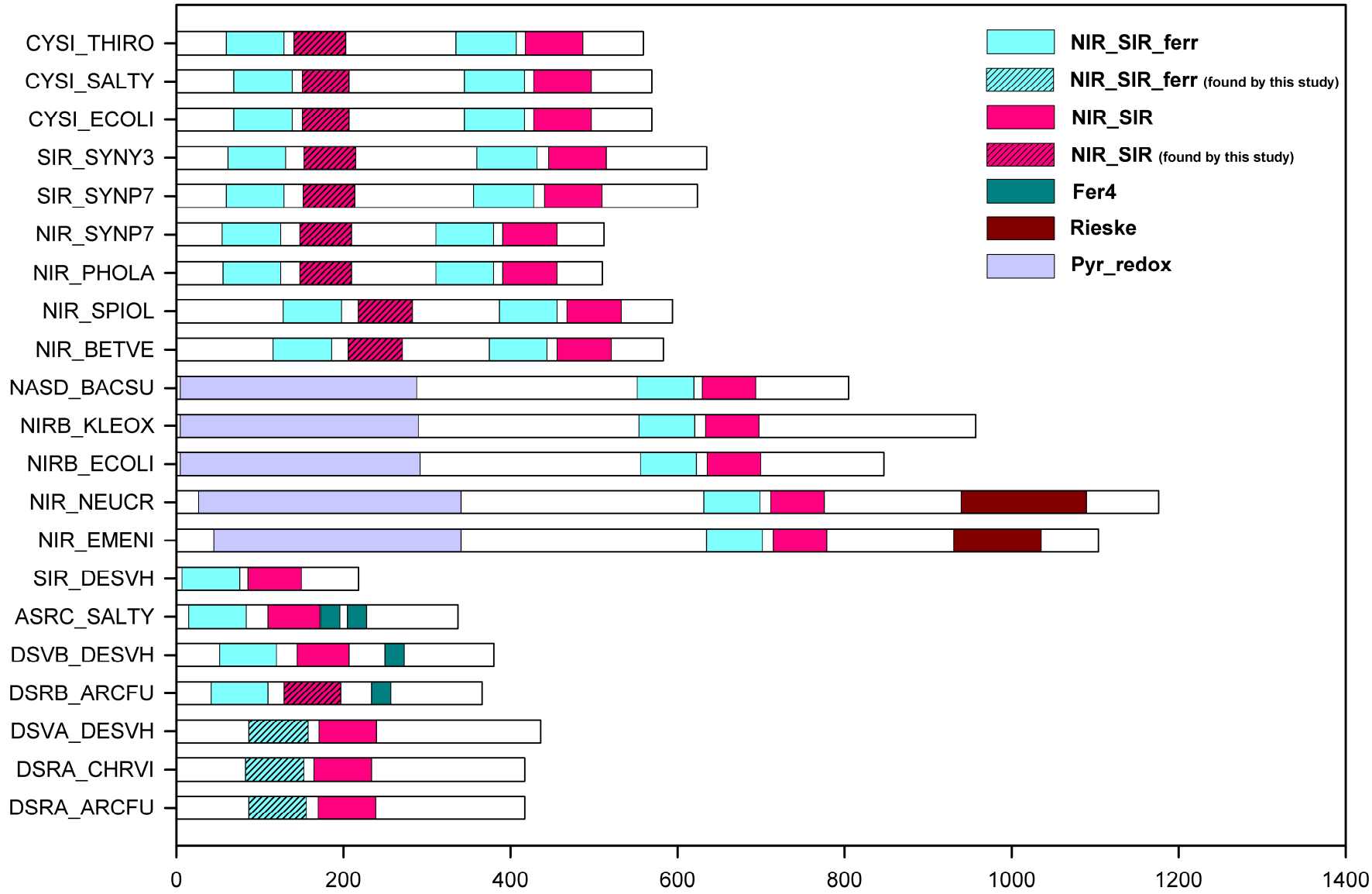


Fig. 2.5: Domain structures of group 2 proteins. SWISS-PROT IDs were used for protein names. Domains are defined following the Pfam database. Hatched domains were found by this study.

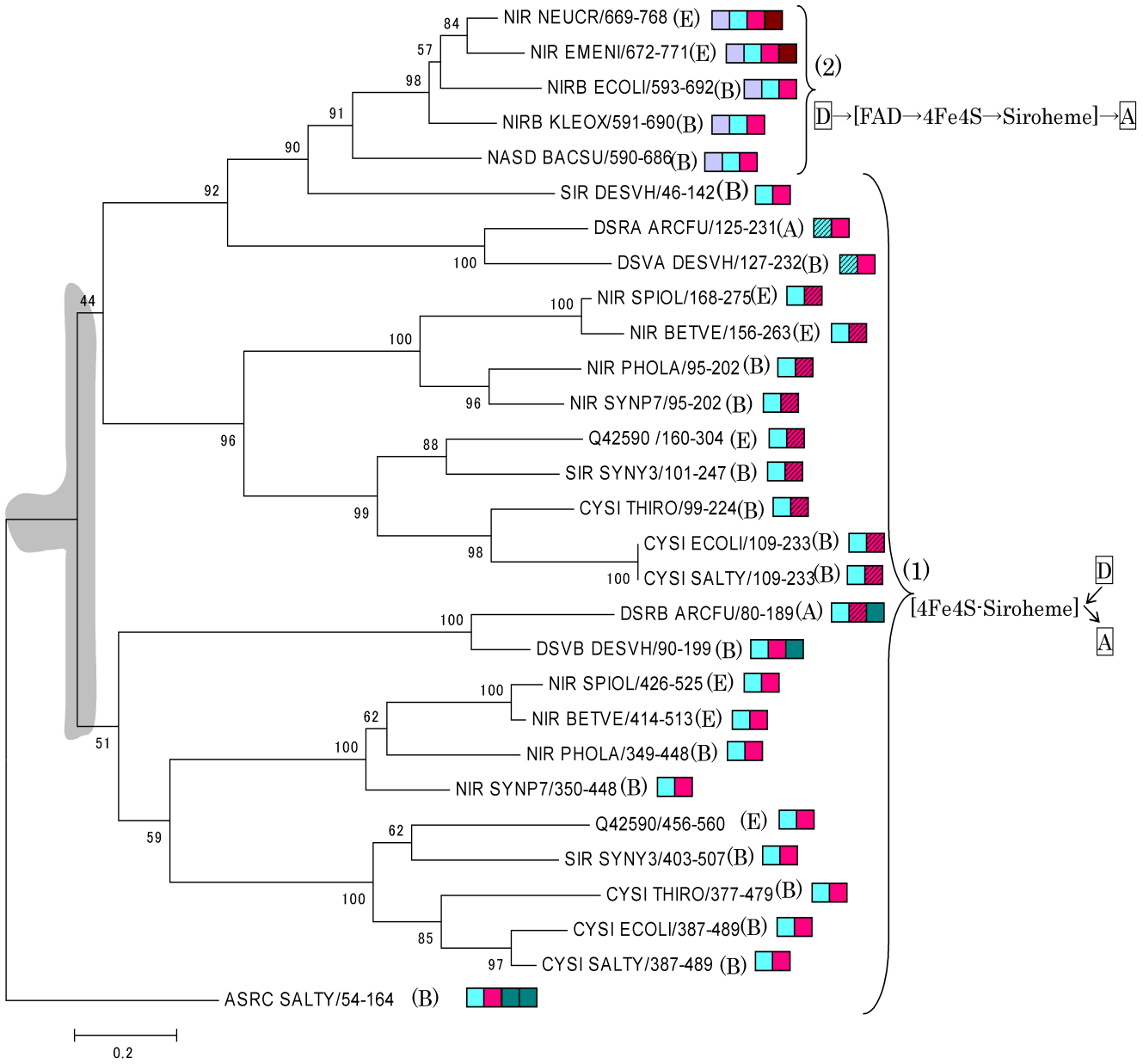


Fig. 2.6: A phylogenetic tree of group 2 proteins based on conserved regions consisting of NIR_SIR_ferr domain and NIR_SIR domain. Conserved regions which are used for constructing this phylogenetic tree were referred to previous works (Crane et al. 1997; Crane and Getzoff 1996; Larsen et al. 1999). Amino acid sequences of sulfite reductase (Q42590) in previous work are also added. The gray colored region shows possible locations of root for this tree. Each colored square in this figure shows domain. Explanatory note of each domain is shown in Fig. 2.5. Arrows of electron transfer patterns indicate the direction of electron flow. D and A surrounded by square mean electron donor and acceptor, respectively. Unknown patterns of electron transfer were predicted from known patterns of evolutionarily closely related proteins. Amino acid sequences of eubacteria, archaea and eukaryotes are designated by (B), (A) and (E), respectively.

Enzymatic reactions of group 2 proteins are accompanied by electron transfer. Arrows show electron flows. Group 2 proteins can be categorized into two types by the differences of electron transfer patterns (Fig. 2.6). Type 1 proteins bind siroheme and 4Fe4S ferredoxin, and electrons from donor are transferred to acceptor at these molecules (Crane et al. 1997; Tan and Cowan 1991). Type 2 proteins bind siroheme and 4Fe4S ferredoxin. However, electron flow is probably different from type 1 as follows; electron donor => FAD => 4Fe4S => siroheme => electron acceptor or substrate (Vega 1976). The tree topology of the phylogenetic tree shown in Fig. 2.6 suggests that the electron transfer pattern for type 1 is ancestral and that for type 2 is derived.

Larsen et al. (1999) noticed that dissimilatory sulfite reductase alpha subunit (DSAR_ARCFU) and beta subunit (DSRB_ARCFU) of *Archaeoglobus fulgidus* may have been duplicated from an ancestral *dsr* gene before the eubacteria and archaea divergence. The tree topology of Fig. 2.6 indicates that this gene duplication was followed by emergence of proteins having a pair of NIR_SIR and NIR_SIR_ferr domains. This suggests that duplicated genes were fused in the first 8 proteins in Fig. 2.5. The root of the tree shown in Fig. 2.6 must be placed before this gene duplication. The gray colored region in Fig. 2.6 thus shows possible locations of root for this tree.

2.3.1.3 Group 3

Homologous group 3 contains 122 proteins, and they were found by using 7 query sequences. Conserved region shared by all the proteins in this group was NAD_binding_1 domain defined by Pfam database. NAD_binding_1 domain has NAD(P)H connection sites. NAD(P)H works as electron donor or acceptor of group 3 protein. Because some proteins were closely related and formed clear monophyletic clusters in the phylogenetic tree when I used

all proteins, I chose one amino acid sequence from each cluster. When more than one enzyme were found in one cluster, I chose representative proteins from each enzyme. A total of 48 proteins were thus used in the following analyses.

Domain structures of group 3 proteins are shown in Fig. 2.7, and they are all flavin enzymes which bind flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN). FAD is bound to FAD_binding_6 domain or FAD_binding_1 domain, and FMN is bound to Flavodoxin or FAD_binding_6 domain.

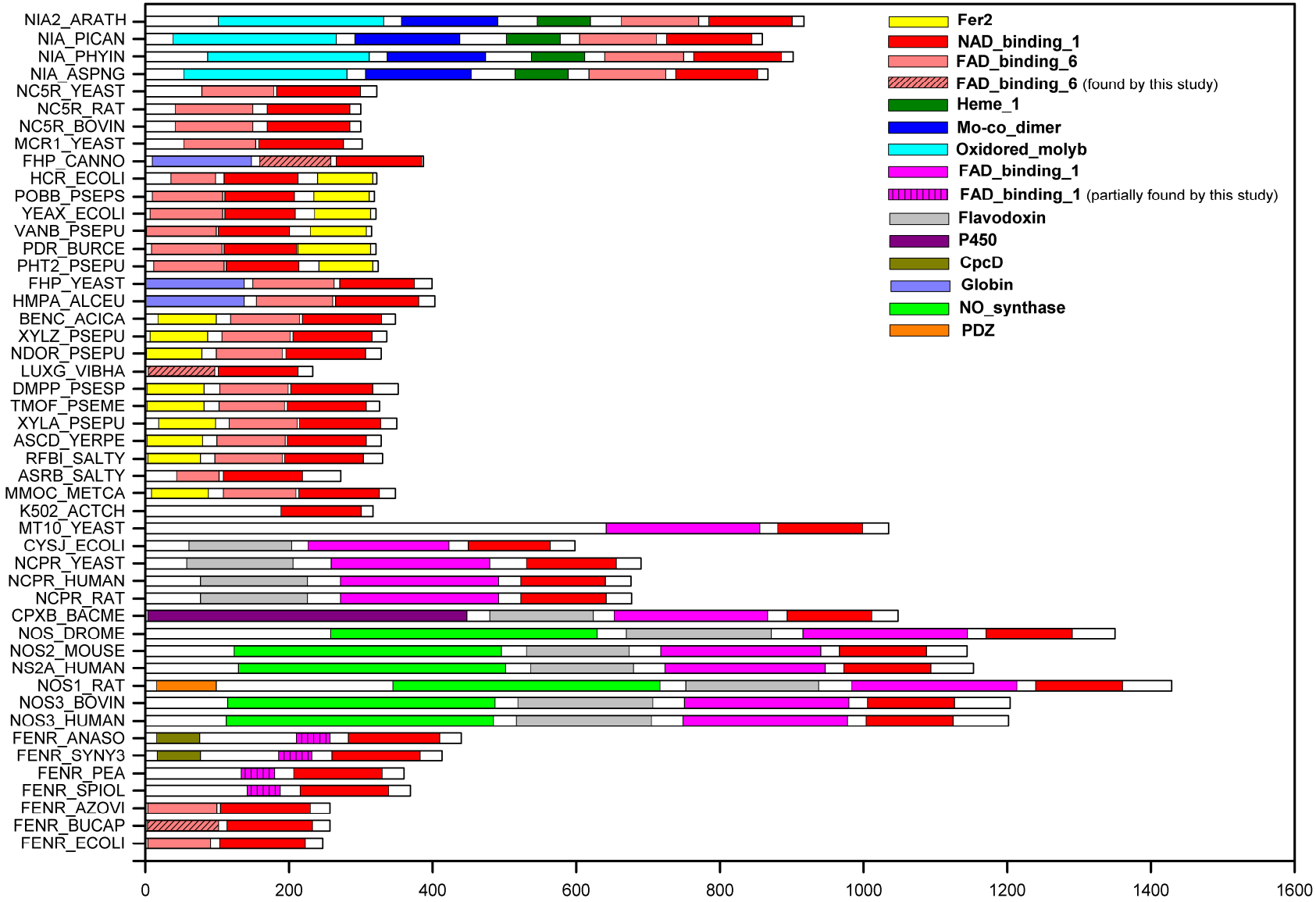


Fig. 2.7: Domain structures of group 3 proteins. SWISS-PROT IDs were used for protein names. Domains are defined following the Pfam database.

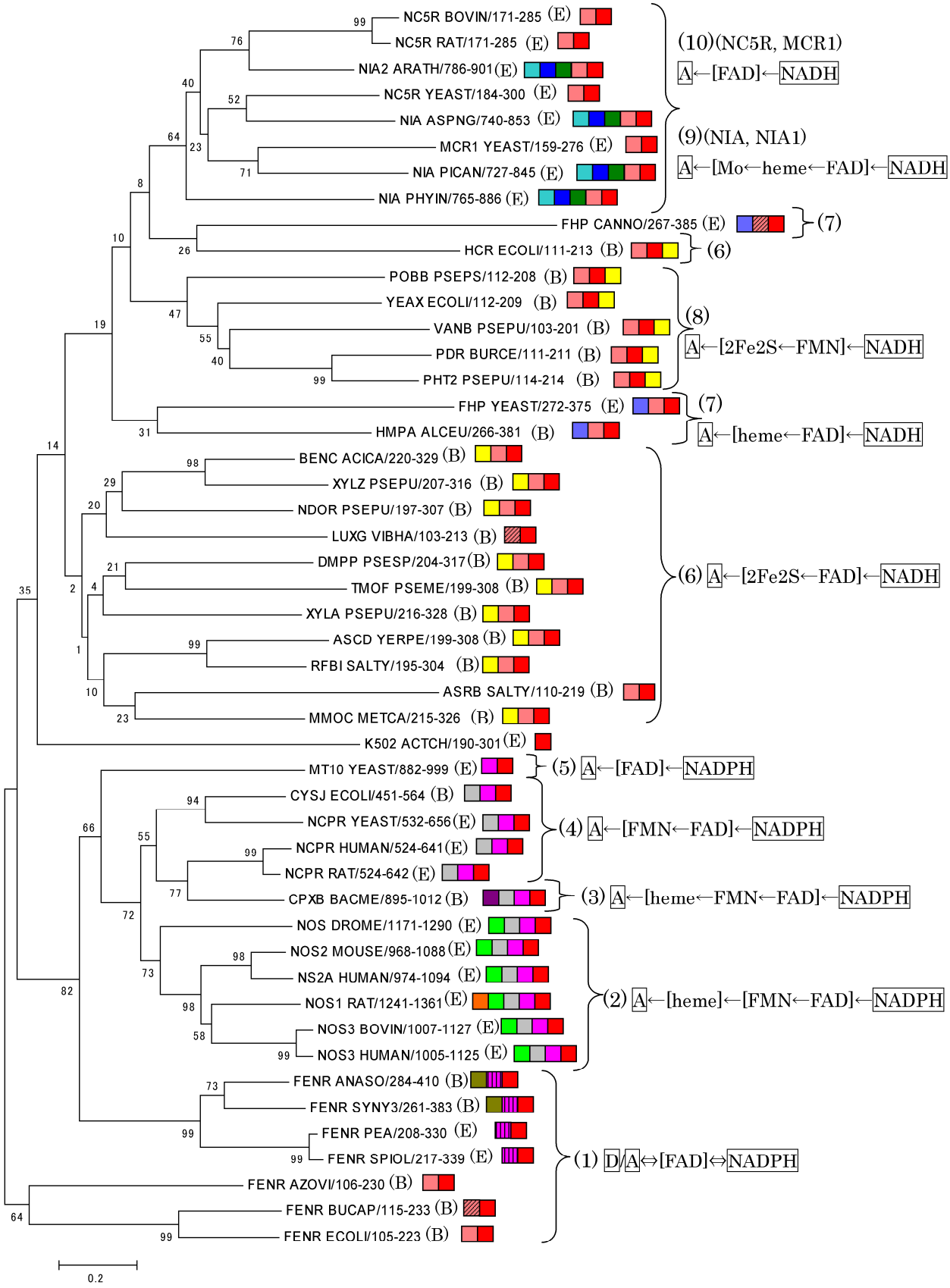


Fig. 2.8: A phylogenetic tree of group 3 proteins constructed based on NAD_binding_1 domain sequences. Each colored square shows domain. Explanatory note of each domain is shown in Fig. 2.7. Arrows of electron transfer patterns indicate the direction of electron flow. D and A surrounded by square mean electron donor and acceptor respectively. Unknown patterns of electron transfer were predicted from known patterns of evolutionarily closely related proteins. Amino acid sequences of eubacteria, archaea and eukaryotes are designated by (B), (A) and (E), respectively.

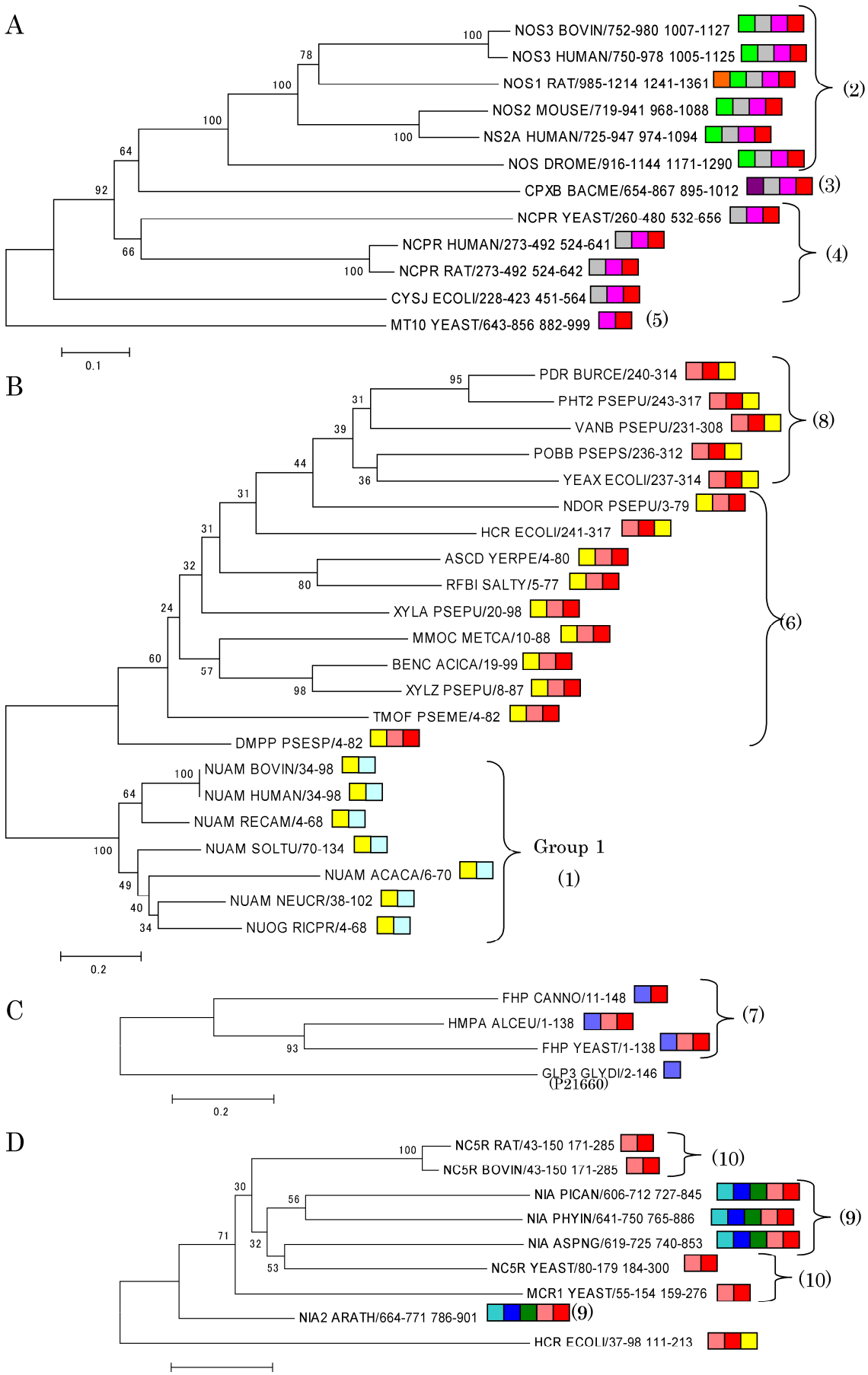


Fig. 2.9: **A.** A phylogenetic tree of the type 2, 3, 4 and 5 proteins based on FAD_binding_1 and NAD_binding_1 domains. **B.** A phylogenetic tree of the type 1 proteins in group 1 and the type 6 and 8 proteins in group 3 based Fer2 domain. **C.** A phylogenetic tree of the type 7 proteins constructed based on Globin domain. **D.** A phylogenetic tree of the type 9 and 10 proteins based on FAD_binding_6 and NAD_binding_1 domains.

Enzymatic reactions of group 3 proteins are accompanied by electron transfer. Arrows in Fig. 2.8 are electron flows. Group 3 proteins can be categorized into ten types based on electron transfer patterns (Fig. 2.8). Type 1 proteins (ferredoxin-NADP reductase) have relatively simple domain structures, and both eukaryotes and eubacteria have this type. Furthermore, looking at the tree topology of Fig. 2.8, type 1 proteins are not monophyletic. Therefore, the common ancestor of group 3 proteins could be similar to that of type 1.

Type 2 proteins (nitric-oxide reductase) are homodimers. FAD, FMN and heme are bound to each subunit. One subunit gets electron from NADPH and gives it to heme of the other subunit through FAD and FMN of the first subunit. Heme of the second subunit gives electron to electron acceptor. Electron from NADPH passes through two subunits and goes to electron acceptor. Electron flow in each pair of brackets shows electron flow occurring in each subunit. Although type 3 protein binds FAD, FMN, and heme as in the case of type 2 proteins, their electron transfer pattern is different from that of type 2 proteins. Type 3 protein is not homodimer. Type 3 protein receives electron from NADPH and gives it to electron acceptor through FAD, FMN, and heme. The NO_synthase domain of type 2 proteins and the P450 domain of type 3 protein have heme binding sites. It suggests that the insertion of NO_synthase domain or P450 domain enabled group 3 protein to bind heme. Type 2, 3, and 4 proteins bind FAD and FMN. The FAD_binding_1 domain has FAD binding sites, and the Flavodoxin domain has FMN binding sites. The insertion of Flavodoxin domain may have enabled group 3 protein to bind FMN.

Fig. 2.9A shows the phylogenetic tree of the proteins belonging to types 2-5, constructed based on FAD_binding_6 and NAD_binding_1 domains. The topology of this tree is more reliable than that of the tree constructed based on only NAD_binding_1 domain sequences. Type 5 proteins do not have Flavodoxin domain. Therefore, Fig. 2.9A tree suggests

that the insertion of Flavodoxin occurred on the lineage to type 2, 3 and 4 proteins. Fig. 2.9A also suggests that NO_synthase and P450 domains that have binding sites of heme were inserted on the lineage to type 2 and type 3 proteins, respectively.

Type 6 and type 8 proteins bind 2Fe2S ferredoxin. Ferredoxin is bound to Fer2 domain. It suggests that insertion of Fer2 domain enabled group 3 proteins to bind ferredoxin. Fig. 2.9B shows the phylogenetic tree constructed based on Fer2 domain. Not only type 6 and type 8 proteins in group 3 have Fer2 domain, but also type 1 proteins in group 1 have this domain. I thus used Fer2 domain sequences of the tree type proteins for constructing the tree shown in Fig. 2.9B. Proteins of this tree can be largely categorized into two lineages; type 1 proteins in group 1, and type 6 and type 8 proteins in group 3. The tree topology of group 3 proteins in Fig. 2.9B suggests that the Fer2 domain was inserted to type 6 proteins, and the same domain was further inserted to form type 8 proteins. The inserted location of the Fer2 domain is different between type 6 and 8 proteins except for HCR_ECOLI (Figs. 2.7 and 2.9B). Although the domain structure of HCR_ECOLI is similar to that of type 8 proteins, the electron transfer pattern of HCR_ECOLI is similar to type 6 proteins. HCR_ECOLI is phylogenetically closer to type 6 proteins (see Fig. 2.9B). One possibility is that domain inversion occurred in the HCR_ECOLI lineage.

Hemes in type 7 proteins are bound to the Globin domain (see Fig. 2.8). The insertion of the Globin domain may have enabled group 3 proteins to bind heme. Type 7 proteins are not monophyletic; HMPA_ALCEU (bacteria) and FHP_YEAST (eukaryote) formed a weak monophyletic cluster (bootstrap probability is only 30%), while FHP_CANNO is located in a different lineage (Fig. 2.8). Fig. 2.9C shows the phylogenetic tree of type 7 proteins constructed based on the Globin domain sequences. This tree also suggests that HMPA_ALCEU and FHP_YEAST form a monophyletic cluster. The bootstrap probability of

this cluster is now 93%. Therefore, it is possible that the lineage of HMPA_ALCEU and FHP_YEAST is separated from that of FHP_YEAST. It implies that the Globin domain may have been inserted independently on the lineage to HMPA_ALCEU and FHP_YEAST and the lineage to FHP_YEAST.

Type 9 proteins are nitrate reductase, and they bind FAD according to the tertiary structure analysis (Lu et al. 1994; Lu et al. 1995). It was hypothesized by sequence similarity that heme and molybdenum are bound to Heme_1 domain and Oxidored_molyb domain of type 9 protein, respectively (Avila et al. 1995; Crawford et al. 1988; Pieterse et al. 1995; Unkles et al. 1992). It suggests that insertion of the Oxidored_molyb domain enabled group 3 proteins to bind molybdenum, and insertion of the Heme_1 domain enabled group 3 proteins to bind heme. Fig. 2.9D shows the phylogenetic tree of type 9 and type 10 proteins constructed based on NAD_binding_1 and FAD_binding_6 domains. The topology of this tree is more reliable than that of the tree of Fig. 2.8 that was constructed based on only NAD_binding_1 domain sequences. This tree suggests that type 9 and 10 proteins are not monophyletic. However, bootstrap probabilities of various branches are not so high, so it is not clear.

2.4 Discussion

2.4.1 Domain Change and Electron Transfer Pattern Change through Evolution

The phylogenetic analysis of the three groups suggests that each homologous group experienced frequent insertions and deletions of domains. I thus estimated domain insertion/deletion events during evolution and connected those events with electron transfer pattern changes.

Group 1. I hypothesized the evolutionary process of group 1 proteins, as depicted in Fig. 2.10A, based on the phylogenetic tree shown in Fig. 2.4. I assumed that type 2 was the ancestral type of electron transfer pattern, and the domain structure of the ancestral protein was assumed to be the same as that of type 2 proteins. Type 2 proteins have Molybdop_Fe4S4 and Molybdopterin domains. Two of type 2 proteins have Molydop_binding domain which is important for binding molybdenum, but FDHA_METJA does not have this domain. As I described above, I suspected that FDHA_METJA does not bind molybdenum and lost the catalytic activity. Because some of NADH dehydrogenase subunits (type 1 proteins) have Molydop_biding domain (Fig. 2.3), it seems that the ancestral protein also had Molydop_binding domain. Therefore, I assumed that the Molydop_binding domain was lost independently on the lineage to FDHA_METJA and on the type 1 proteins (Fig. 2.10). If this hypothesis is true, the ancestral protein probably had Molydop_Fe4S4, Molybdopterin, and Molydop_binding domains as in the case for the type 2 proteins except for FDHA_METJA. The ancestral protein may have bound molybdenum and 4Fe4S ferredoxin.

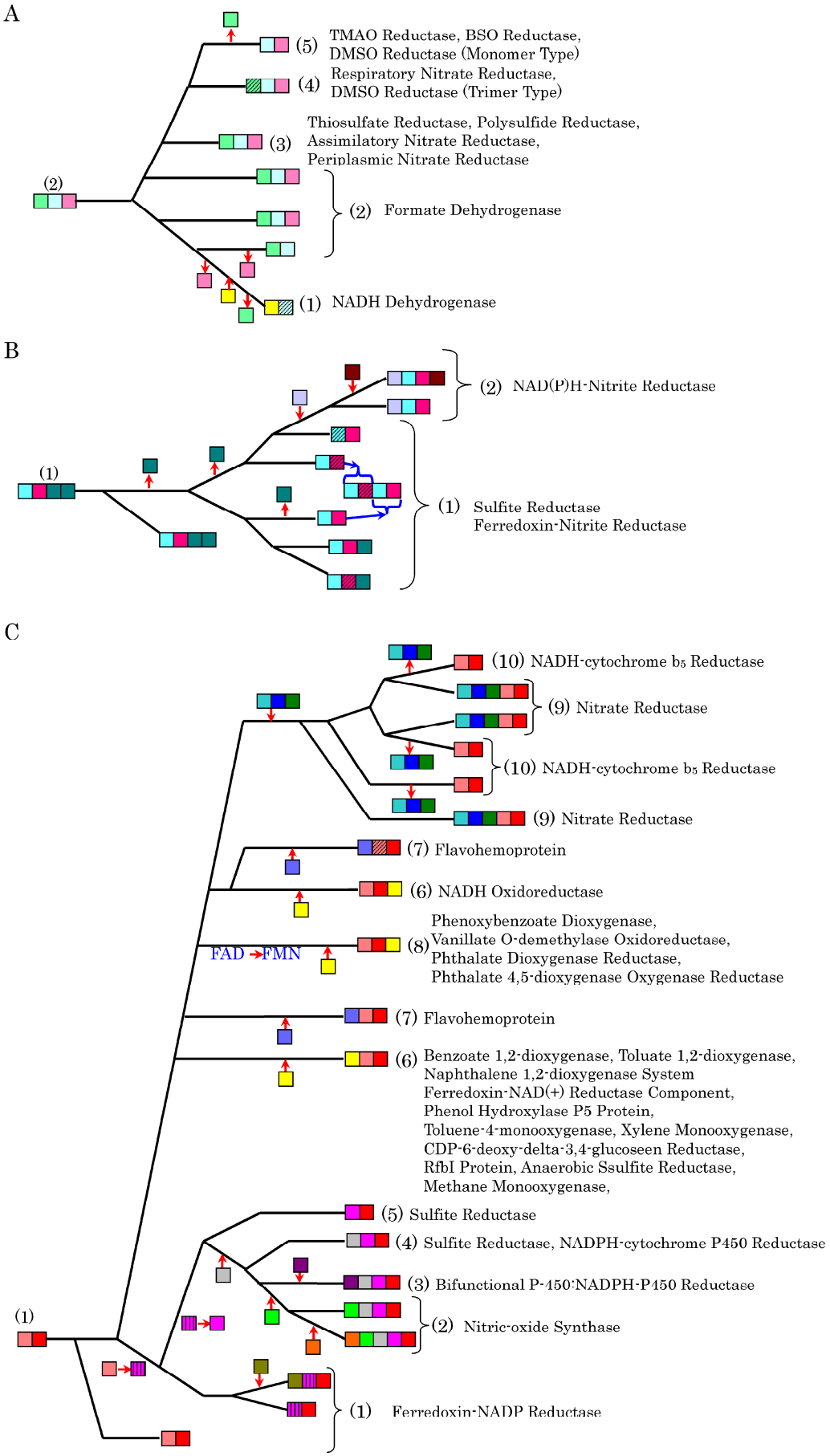


Fig. 2.10: Estimated domain insertion/deletion events during evolution of the three protein groups. Each square shows domain defined by Pfam. **A.** Group 1 proteins. Color patterns are same as in Fig. 2.3. **B.** Group 2 proteins. The correspondence between color and Pfam domain is shown in Fig. 2.5. **C.** Group 3 proteins. Color of each domain is same as that in Fig. 2.7.

A scenario of evolutionary changes of the group 1 proteins from this ancestral protein is as follows. On the lineage to type 1 proteins (NADH dehydrogenase), Fer2 domain was acquired, and ferredoxin was bound to this domain. In contract, the Molybdopterin domain became less conserved, and most of type 1 proteins lost Molydop_binding domain. Both Fer2 and ferredoxin domains are important for molybdenum binding. Therefore type 1 proteins lost the function of binding molybdenum which is important for catalysis. Its main function changed from catalysis to electron transfer. On the lineage to type 3, the direction of electron flow was reversed. The molybdop_4Fe4S domain that has binding sites of 4Fe4S ferredoxin became less conserved on the lineage to type 4 proteins, and was completely lost on the lineage to type 5. Types 4 and 5 proteins lost the function of binding 4Fe4S ferredoxin. In type 5 proteins, electron is exchanged at molybdenum from electron donor to electron acceptor. There are no other electron carriers in this protein.

The common feature of the group 1 proteins is that all of them have Molybdopterin domain (Fig. 2.3), although this domain is not conserved well in type 1 proteins. Other features are not necessarily common. For example, ferredoxin exists in types 1-3, but not in others (Fig. 2.4). Molybdenum exists in types 2-5, but not in type 1 (Fig. 2.4). Some of group 1 proteins (with * in Fig. 2.4) are subunits of proteins which can generate proton gradient. Proton gradient is the energy source for generating ATP by ATP synthase. From my phylogenetic analysis, this important function, proton gradient generation, is not necessarily conserved. Although the general electron transfer patterns are more or less similar in type 3 proteins, some proteins (polysulfide reductase subunit and thiosulfate reductase subunit) work for proton gradient generation, but the others (assimilatory nitrate reductase subunit) do not (see Fig. 2.4).

Group 2. I conjectured the evolution of group 2 proteins as shown in Fig. 2.10B. The common ancestral protein was assumed to have both the NIR_SIR_ferr and NIR_SIR domains, because all the group 2 proteins have these two domains (including weakly conserved ones). The two Fer4 domains were assumed to exist in the common ancestral protein, as in the case of ASRC_SALTY protein. Under this assumption, three losses of Fer4 domain were assumed (Fig. 2.10B). However, proteins with only one or no Fer4 domain can also be common ancestor. All three possibilities are equally parsimonious for three insertion/deletion events of Fer4 domain.

Later, a gene duplication produced two copies of NIR_SIR_ferr and NIR_SIR domains. It means that one gene including NIR_SIR_ferr and NIR_SIR domains was duplicated, and two genes like a pair of DSRA (alpha subunit of sulfite reductase) and DSRB (beta subunit) or that of DSVA (alpha subunit of sulfite reductase) and DSVB (beta subunit) were produced. These two genes were later fused. In the lineage of DSRA and DSVA, Pyr_redox domain was added and FAD binding site was generated. This FAD binding seems to be the critical point for generating the type 2 proteins.

Group 3. A parsimoniously hypothesized evolutionary pathway of group 3 proteins is shown in Fig. 2.10C. As I described in the Results section, the electron transfer pattern of the ancestral proteins seems to have been type 1. If so, that ancestral protein had NAD_binding_1 and FAD_binding_6 domains like type 1 proteins of eubacteria. On the lineage to types 2, 3 and 4, Flavodoxin domain having the FMN binding region was inserted. Because of this addition, FMN joined as a member of electron flow. Domains of P450 and NO_synthase that have binding sites for heme were inserted on the lineages to type 2 and 3 proteins, respectively. Type 2 and type 3 proteins thus acquired the ability to bind heme. On the lineages to types 6 and 8, 2Fe2S ferredoxin was added by the insertion of Fer2 domain which

has a 2Fe2S ferredoxin binding sites. FAD was replaced to FMN in type 8, although the FAD_binding_6 domain that has the binding region of FAD or FMN did not change. Heme was also added on the lineage to type 7 proteins because of being inserted Globin domain which has heme binding sites. Lineages to types 9 and 10 proteins are a little complicated. Based on the phylogenetic tree in Fig. 2.9D, I hypothesized that the insertion of the three domains (Oxidored_molyb, Mo-co_dimer and Heme_1) had occurred once. The three domains were then deleted three times on the lineage to MCR1_YEAST, the lineage to NC5R_YEAST and the lineage to NC5R_BOVIN and NC5R_RAT independently. But the branching pattern of types 9 and 10 is not supported by high bootstrap probabilities. Therefore, I can not predict how these proteins were evolved in detail. In Fig. 2.10C, I illustrated the lineages of type 9 and 10 proteins which I hypothesized based on Fig. 2.9D.

The common features of the group 3 proteins are that all proteins have NAD_binding_1 domain. NAD(P)H is their electron donor or acceptor, and FAD or FMN is bound to them as an electron transfer component. Although these features are conserved, electron transfer patterns in group 3 proteins are so diverged by adding and deleting various domains.

2.4.2 Evolution of the Energy Generation Metabolisms and the Assimilatory Metabolisms

The four energy metabolisms and dissimilation in sulfur metabolism are energy generation metabolisms. Proteins which can generate proton potential (abbreviations of such proteins are red-colored in Fig. 2.11) are conserved between different metabolism systems. Proton potential is the source for the synthesis of ATP by ATP synthase. This function is important for energy generation metabolisms. Proteins which have this important function,

generation of proton potential, seem to be conserved among the four energy metabolisms and dissimilation in sulfur metabolism. This suggests that those systems are evolutionarily related.

I therefore tried to infer the evolutionary relationship of the four energy metabolisms and dissimilation metabolism. Trees A-E of Fig. 2.12 show the five phylogenetic trees of homologous proteins analyzed in this study and previous analyses (Berry et al. 2000; Saraste et al. 1994). Divergence between eubacteria and others (archaea and eukaryotes) are shown in Fig. 2.12 based on rough estimation of this study (Figs. 2.4, 2.6, and 2.8) and previous studies (Berry et al. 2000; Saraste and Castresana 1994).

Those protein trees were superimposed to infer the evolutionary tree of metabolic systems. The idea of superimposition came from OOta and Saitou (1999), who inferred the evolutionary tree of muscle tissues by superimposition of muscle protein trees. Because the protein cluster of denitrification (N_{DN}), sulfur respiration (S_R), and dissimilation in sulfur metabolism (S_{DS}) was estimated to emerge after the divergence of eubacteria and archaea/eukaryotes according to tree A, these three systems were considered to be closely related. Therefore, N_{DN} of trees B and C of Fig. 2.12 were assumed to correspond to the N_{DN} - S_R - S_{DS} cluster of tree A. Trees A, B, and C of Fig. 2.12 were thus superimposed to produce tree F.

I then superimposed trees D and E of Fig. 2.12. Because my interest is the order of system generation, I ignored the newer lineage of nitrate assimilation (N_{AS}) and assimilation in sulfur metabolism (S_{AS}) in these two trees. The resultant superimposed tree is tree G of Fig. 2.12. Trees F and G were further superimposed to produce single phylogenetic tree H of the seven metabolic systems. However, because of the lack of information, the branching point of the aerobic respiration system was not able to be determined, and there

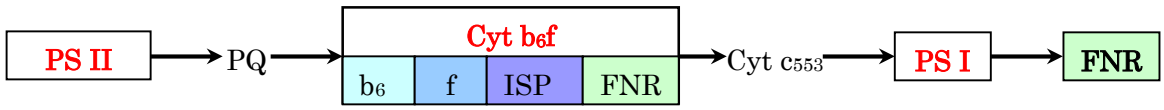
are five branching possibilities (①-⑤ in tree H of Fig. 2.12). Tree H suggests that photosynthesis is basal (diverged first) among the energy metabolisms I analyzed.

If I accept tree H, the assimilation process (nitrate assimilation and assimilation in sulfur metabolism) seems to be older than the dissimilation process (denitrification, sulfur respiration and dissimilation in sulfur metabolism). Assimilation systems might be generated before the divergence of eubacteria and archaea/eukaryotes, and dissimilation systems might be generated after the divergence of eubacteria and archaea/eukaryotes.

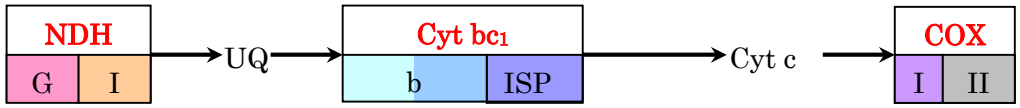
Although tree H does not tell the ancestral system at the branching nodes, I can infer that nitrate assimilation was ancestral to the sulfur assimilation metabolism. Nitrate assimilation is involved in the metabolism for producing any amino acids, while assimilation in sulfur metabolism is involved only for producing amino acids including sulfur (cysteine and methionine). It is conceivable that generation of nitrate assimilation enabled ancient organisms to produce amino acids except for cysteine and methionine at first, followed by acquisition of the ability for producing cysteine and methionine by assimilation in sulfur metabolism.

Fig. 2.11 (next page): Homologous proteins in the four metabolisms (photosynthesis, oxidative phosphorylation, nitrogen metabolism and sulfur metabolism). Black and blue arrows designate electron flow and compound change (enzymatic reaction), respectively. Square objects are a proteins. When one protein is composed by subunits, small squares showing subunits are shown in the lower part of the square. Some proteins involved in the four metabolisms are not shown in Fig. 2.11. Same colored proteins are homologous. The explanatory notes of colors are shown at the bottom of this figure. Each color shows each homologous group. Purple red colored proteins (COX1) were not detected by homology search analysis in this study, but they are thought to be homologous (Mogi et al. 1998; Saraste and Castresana 1994; Zumft et al. 1992). Subunits A to D and H to N of NDH are put together. Same alphabetical subunits of NDH in photosynthesis and NDH in oxidative phosphorylation are homologous. Conserved Pfam domains in each homologous group are shown in parenthesis of explanatory notes. Abbreviations in this figure are as follows. PS II = photosystem II; NDH = NADH dehydrogenase; PQ = plastoquinone; Cyt b_6f = cytochrome b_6f complex; ISP = iron sulfur protein (subunit); FNR = ferredoxin NADP reductase; Cyt c_{553} = cytochrome c_{553}; PS I = photosystem I; UQ = ubiquinone; Cyt bc_1 = cytochrome bc_1 complex; Cyt c = cytochrome c; COX = cytochrome c oxidase; NAR = respiratory nitrate reductase; NAP = periplasmic nitrate reductase; NAS = assimilatory nitrate reductase; Cat = catalytic subunit; Cyt c_{550} = cytochrome c_{550}; NIA = eukaryotic assimilatory nitrate reductase; NIR = nitrite reductase; NOR = nitric oxide reductase; NosZ = nitrous oxide reductase; MQ = menaquinone; PSR = polysulfide reductase PHS = thiosulfate reductase; DSR = sulfite reductase dissimilatory type; SIR = sulfite reductase (ferredoxin); CYS = sulfite reductase (NADPH); ASR = anaerobic sulfite reductase

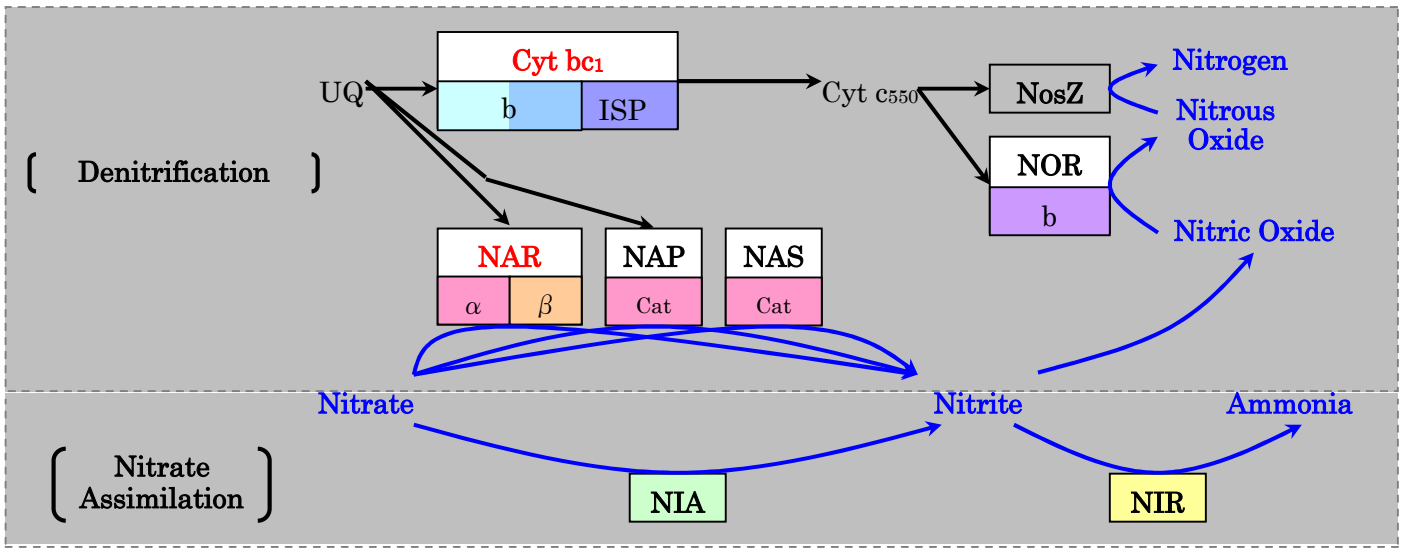
Photosynthesis



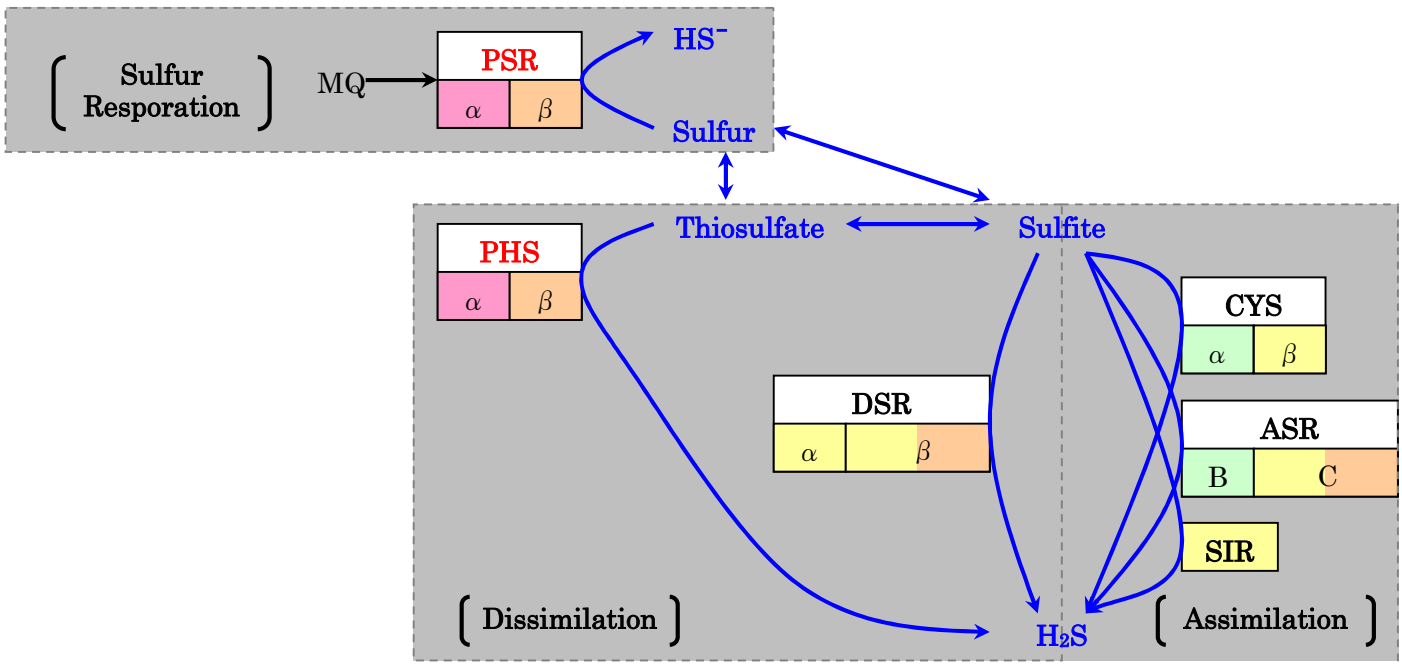
Aerobic Respiration



Nitrogen Metabolism



Sulfur Metabolism



■: Group 1 (Molybdopterin) ■: Group 2 (NIR_SIR, NIR_SIR_ferr) ■: Group 3 (NAD_binding_1)
■: Group 4 (Fer4) ■: Group 5 (Cytochrome_b_N) ■: Group 6 (Cytochrome_b_C) ■: Group 7 (Rieske)
■: Group 8 (not defined by Pfam) ■: Undetected (COX1)

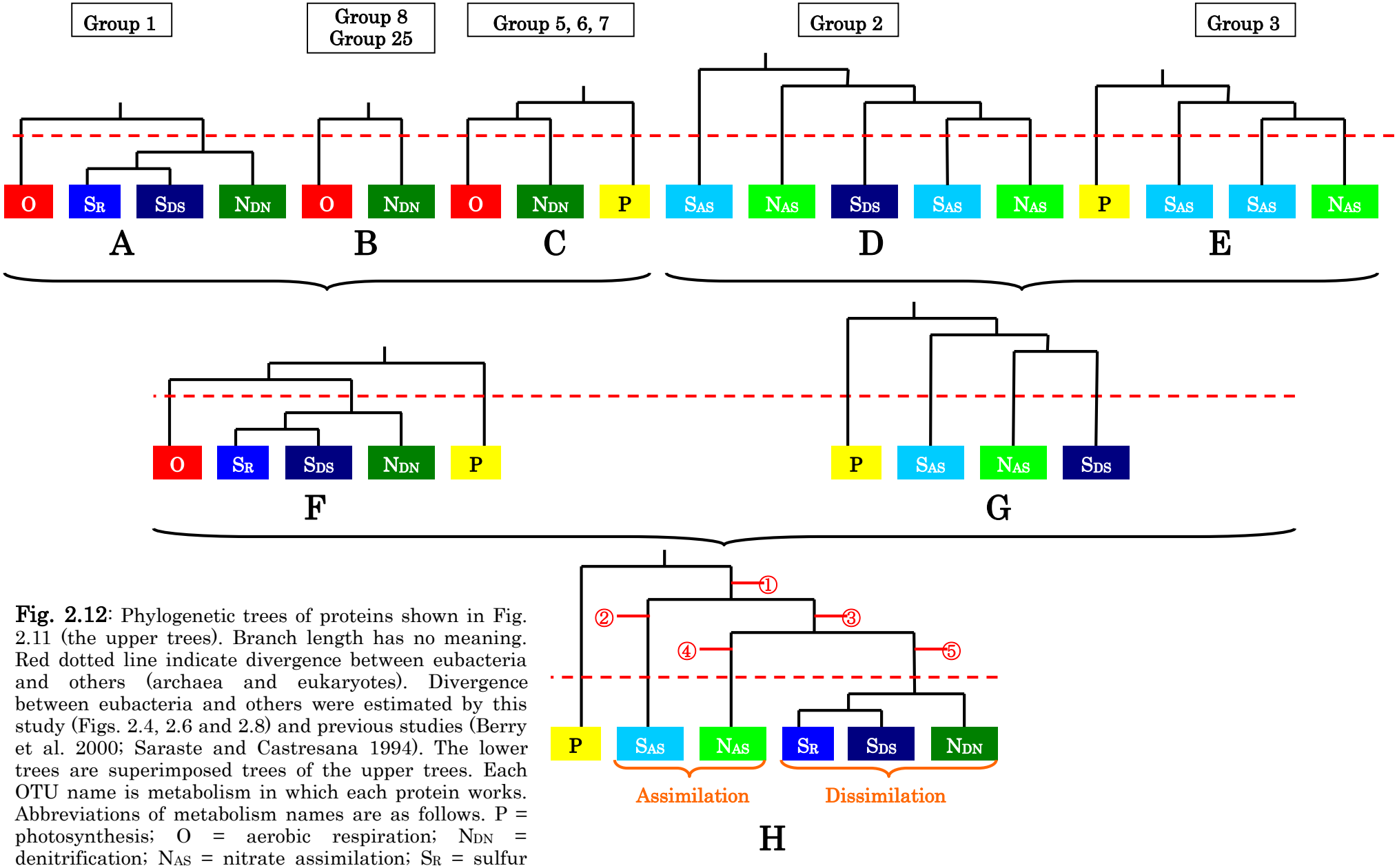


Fig. 2.12: Phylogenetic trees of proteins shown in Fig. 2.11 (the upper trees). Branch length has no meaning. Red dotted line indicate divergence between eubacteria and others (archaea and eukaryotes). Divergence between eubacteria and others were estimated by this study (Figs. 2.4, 2.6 and 2.8) and previous studies (Berry et al. 2000; Saraste and Castresana 1994). The lower trees are superimposed trees of the upper trees. Each OTU name is metabolism in which each protein works. Abbreviations of metabolism names are as follows. P = photosynthesis; O = aerobic respiration; N_{DN} = denitrification; N_{AS} = nitrate assimilation; S_R = sulfur respiration; S_{DS} = dissimilation in sulfur metabolism; S_{AS} = assimilation in sulfur metabolism

As I have seen in this study, insertions and deletions of domains produce diverse patterns of electron transfer which is essential for enzymatic function. But the diversity of electron donor and acceptor is not produced by insertions and deletions of domains. A small number of amino acid substitutions may produce the diversity of electron donor and acceptor. However, I did not conduct the analysis of a small number of amino acid substitutions because of scanty experimental data. Homologous domains are found not only within the energy generation metabolisms but also both in the energy metabolisms and the assimilatory metabolisms. The functions of these homologous domains are involved in electron transfer. This type of function can be used widely in various metabolisms. Combinations of functional domains produce diversity of protein function and functional domains can be used repeatedly in various metabolisms through evolution.

3. Phylogenetic Analysis of Proteins in Neurotransmission Systems

3.1 Introduction

Communications between neurons are realized through chemical transmission or electrical transmission. Neurotransmitters as neurotransmission signals between neurons are transferred from presynaptic cells to postsynaptic cells in chemical neurotransmission. Three molecules are important for chemical neurotransmission except for peptides and some neurotransmitters. They are synthases, receptors and transporters of neurotransmitters (Fig. 3.1). Synthases produce neurotransmitters and neurotransmitters are released from presynaptic terminals when presynaptic neurons are activated. The released neurotransmitters bind to receptors of postsynaptic cells, and the chemical signals from the presynaptic neurons are transmitted to the postsynaptic neurons. Neurotransmitters in synapses are uptaken by transporters of the presynaptic neurons.

How have the chemical neurotransmission system composed of the three classes of molecules evolved? Have these three important molecules evolved together or not? If these three important molecules have evolved together, the evolution of the chemical transmission system can be hypothesized as follows. The three important molecules were generated at first, and chemical neurotransmission system was established. After that various neurotransmitters gradually became used. But if the three important molecules have not evolved together, the evolution of the chemical neurotransmission system is as follows. Each molecule had evolved independently, and these three molecules started to work together by chance. I did phylogenetic analysis of

the three important molecules to confirm which hypothesis is correct.

I am also interested in how these three important molecules have evolved. For example, ligand specificities of each molecule have changed by amino acid substitutions through evolution. I analyzed the relationship between amino acid sequences and functions of the molecules, and predicted the amino acid substitutions which are critical for functional changes based on the results of the phylogenetic analysis.

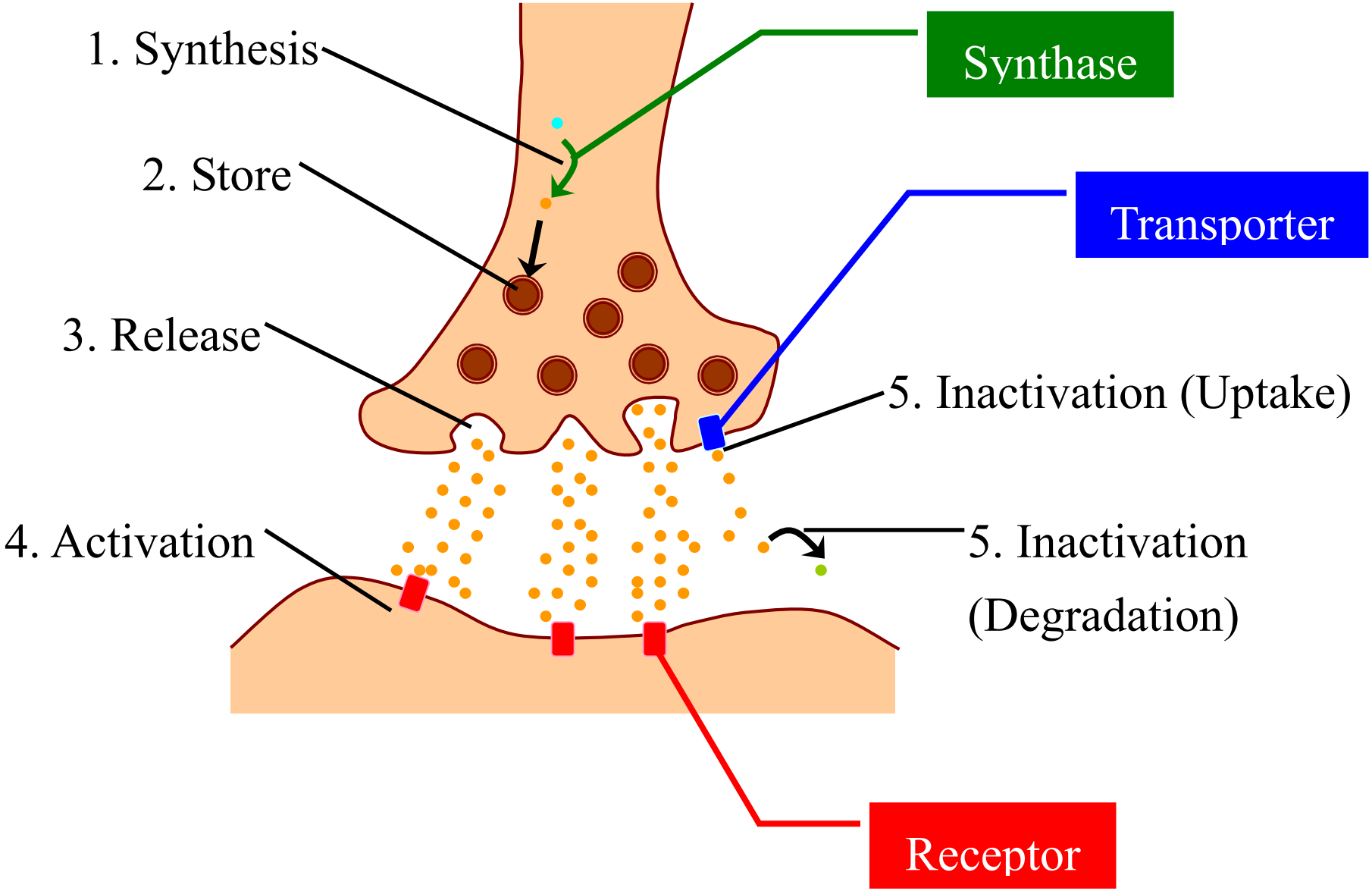


Fig. 3.1: Chemical neurotransmission. Neurotransmissions are synthesized by synthases (1), and stored into vesicles (2). When presynaptic neuron is activated, neurotransmitters are released to synapse (3). Neurotransmitters in synapse are captured by receptors, and postsynaptic neuron is activated (4). Neurotransmitters in synapse are uptaken by transporters or degradated by enzymes, and chemical neurotransmission is inactivated (5).

3.2 Materials and Methods

3.2.1 Amino Acid Sequence Retrieval and Categorization

I retrieved the amino acid sequences of the synthases, the receptors and the transporters of the neurotransmitters listed in Alexander and Peters (2000) and NeuronDB (http://senselab.med.yale.edu/senselab/NeuronDB/) from SWISS-PROT and TrEMBL (Boeckmann et al. 2003) except for neuropeptides whose synthases and transporters do not exist. I used these sequence data as queries and searched SWISS-PROT and TrEMBL by using BLAST2 (Altschul et al. 1990; Altschul et al. 1997; Tatusova and Madden 1999). Based on the results, I categorized the enzymes, the receptors and the transporters into groups by sequence homology.

3.2.2 Database Construction

Amino acid sequence data in each homologous group were integrated into web-based SQL database (PostgreSQL) and I named it "Macaroni" (http://neuron.genes.nig.ac.jp/macaroni/), for transmembrane regions of receptors and transporters look like macaroni. This database has five tables, and each table has relationships to other tables (lines between tables in Fig. 3.2). To entry data, administrator (myself) of the Macaroni database inputs only two data; accession number and category data defined by the administrator. Categories in the database are mainly based on the homologous groups. First major categories are synthases, receptors and transporters (Top Menu in Fig. 3.3). Each protein data is linked to SWISS-PROT or TrEMBL on GenomeNet (http://www.genome.ad.jp) (① in Fig. 3.3). Amino acid sequence data of each protein is also linked to SWISS-PROT or TrEMBL on GenomeNet (② in Fig. 3.3). If amino acid sequence data is fragment, red "F" is shown (② in Fig.

3.3). Domain composition of each protein can be easily to known by Pfam through direct link (③ in Fig. 3.3). Users can easily retrieve amino acid sequence data of proteins in Macaroni. After checking boxes of proteins of which one wants to get sequences (④) and clicking the "get the sequences" button (⑤), users can retrieve sequences of checked proteins. The sequences can store in virtual basket of the Macaroni site (⑥), and users can convert to some sequence formats, or do multiple alignments and construct phylogenetic tree by ClustalW on GenomeNet or DDBJ.

T_INDEX

Field Name	ID †	DB*	OS	FRAG	PFAM	PDB	DE
Data Type	VARCHAR(20)	VARCHAR(3)	TEXT	INT2	TEXT	TEXT	TEXT

T_AC

Field Name	NUM †	AC*	ID*
Data Type	VARCHAR(20)	VARCHAR(20)	VARCHAR(20)

T_PROT

Field Name	AC †	CAT_ID*	PMID
Data Type	VARCHAR(20)	VARCHAR(39)	TEXT

T_CAT

Field Name	CAT_ID †	CAT_NAME#*	CAT_DE
Data Type	VARCHAR(39)	VARCHAR(50)	VARCHAR(254)

T_PUBMED

Field Name	PMID †	TITLE*	AUTH*	JN*
Data Type	VARCHAR(10)	VARCHAR(254)	VARCHAR(254)	VARCHAR(254)

†: Primary Key, #: Unique, *: Not Null, Blue Characters: INDEX

Fig. 3.2: Tables in Macaroni database. Abbrebiations of the field name in each field are as follows. ID = ID number of each sequence; DB = database where sequences are existed; OS = species name; FRAG = fragment data or not; PFAM = IDs of pfam domains which each protein has; PDB = pdb IDs, if tertiary structure of each protein is available; DE = description of each domain; NUM = non-redundant number for identification of each data; AC = accession number of each protein; CAT_ID = IDs of categories we defined; PMID = PubMed IDs of references in which functions of each protein describe; CAT_NAME = category name; CAT_DE = description of each category; TITLE = title of reference; AUTH = author of reference; JN = journal name of reference

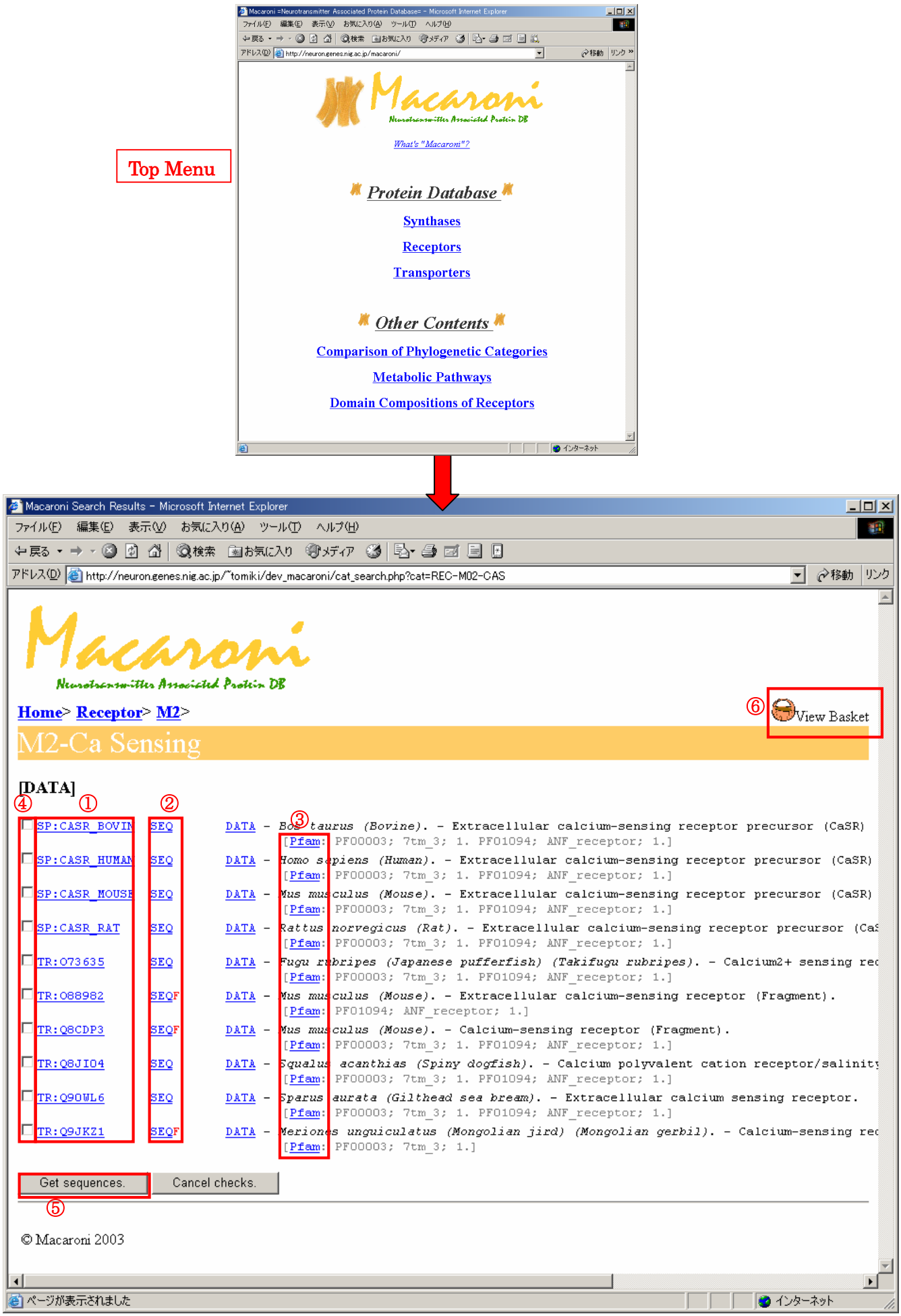


Fig. 3.3: Features of Macaroni

3.2.3 Domain Composition Analysis

I searched Pfam database (http:// www.sanger.ac.uk/Software/Pfam/; Bateman et al. 2002), a collection of protein families and domains, for functional domains of each protein.

3.2.4 Phylogenetic Analysis

Amino acid sequences which I analyzed are so diverged, and proteins in some homologous groups are composed of several domains. Therefore, I constructed phylogenetic trees by using amino acid sequences of each functional domain to obtain reliable trees and know domain composition changes through evolution. Multiple alignment of each domain is available from the Pfam web site. I used multiple alignments available in Pfam database, and constructed neighbor-joining trees (Saitou and Nei 1987) by using MEGA2 (Kumar et al. 2001). Numbers of amino acid substitutions were computed for Poisson correction.

Multiple alignments from Pfam database were not used for the phylogenetic analyses of S1 synthase, soluble guanylyl cyclase and 1.2 receptor. Phylogenetic trees of S1 synthase and soluble guanylyl cyclase constructed by using multiple alignments of the functional domains from Pfam database are not reliable, because the functional domains I used for constructing trees are short. Therefore, I used all sequence regions in S1 synthase and soluble guanylyl cyclase and made the multiple alignments by ClustalX (Thompson et al. 1997). The multiple alignments of the ligand binding domains (S1 and S2 domains) in 1.2 receptors are not available in Pfam database, therefore I made the multiple alignment of the ligand binding domains by ClustalX. The constructions of S1 synthase, soluble guanylyl cyclase and 1.2 receptor were done by the

same way as I described in the last paragraph.

Phylogenetic trees which have same domains were superimposed. Some domains are shared among some homologous groups, and these domains can connect phylogenetic trees constructed by using domains which existed only in each homologous group. I estimated the divergences of species. When I connected phylogenetic trees, I decided the connecting points between phylogenetic trees based on the divergences.

3.2.5 Analysis of Domain Changes and Amino Acid Substitutions

Analyses of the relationship between amino acid sequences and protein functions were applied to receptors. Domain compositions of receptors in some homologous groups were changed through evolution. I analyzed the relationship between domain compositions and receptor functions. Receptor functions were retrieved from annotations in SWISS-PROT and TrEMBL, and bibliographies.

Furthermore, I analyzed the relationship between amino acids of ligand binding sites and ligand binding functions. Ligand binding functions can be changed by several amino acid substitutions. I retrieved amino acid binding sites from bibliographies and parsimoniously predicted amino acid substitutions of each site through evolution based on phylogenetic trees I constructed. Based on the amino acid substitutions I predicted, I hypothesized the essential amino acid substitutions for changing ligand binding functions.

3.3 Results and Discussion

3.3.1 Phylogenetic Analysis of Neurotransmission Systems

I categorized amino acid sequences of enzymes, receptors and transporters of neurotransmitters except for peptides into groups by sequence homology. Table 3.1 is the list of the homologous groups. Chemical structures and metabolic pathways of neurotransmitters are shown in Fig. 3.4. The homologous groups of synthases and transporters were named by myself, and the names of homologous receptor groups were IUPHAR Receptor Code (Humphrey and Barnard 1998). ATP and acetylcholine are degraded to adenosine and choline, and they are uptaken by adenosine transporter and acetylcholine receptor, respectively. The mechanism of neurotransmission by NO is different from others, NO is transmitted from the postsynaptic cell to the presynaptic cell.

Synthases, receptors and transporters of the same homologous groups are used for some neurotransmitters. S1 synthases and 1.1 receptors are used for GABA, histamine and 5-HT. S1 synthases and 2.1 receptors are used for histamine, 5-HT and dopamine. 1.1 receptors and T1 transporters are used for GABA, Glycine and 5-HT. 2.1 receptors and T1 transporters are used for 5-HT, dopamine, adrenaline and noradrenaline. It may be possible to hypothesize that S1 synthases and 2.1 receptors for histamine, dopamine and 5-HT have evolved together, and 2.1 receptors and T1 transporters for 5-HT, dopamine, adrenaline and noradrenaline have evolved together. Other homologous group combinations of synthase and receptor, receptor and transporter, or synthase and transporter are random. Therefore, synthases, receptors and transporters except for S1 synthases, 2.1 receptors and T1 transporters may have evolved independently.

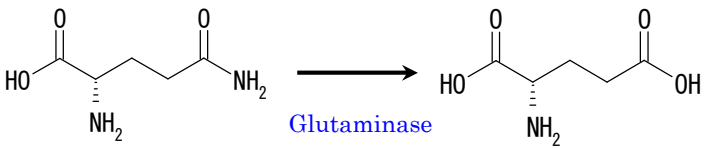
Table3.1: Homologous groups of synthases, receptors and transporters

	Neurotransmitter	Synthase	Receptor*	Transporter
Amino Acid	Glutamate	S5	1.2 / 2.3	T2
	Glycine	S2	1.1	T1
	GABA	S1	1.1 / 2.3	T1
Amine	Histamine	S1	1.1 / 2.1	Degradation
	5-HT (Serotonin)	S1	1.1 / 2.1	T1
	Dopamine	S1	2.1	T1
	Adrenaline	S4	2.1	T1
	Noradrenaline	S3	2.1	T1
Purine	ATP	S6 / S7 / S8	1.4 / 2.1	Degradation (Adenosine Transporter)
	Adenosine	S9	2.1	T3 / T4
Others	Acetylcholine	S10	1.1 / 2.1	Degradation (Choline Transporter)
	Cannabinoid	S12	2.1	?
	Prostanoid	S13 / S14 / S15 / S16 / S17	2.1	T5
	Taurine	S18	1.1?	T1
	NO	S11	Sol-GC	Degradation

* 1.1, 1.2, 1.4: Ionotropic Receptors / 2.1, 2.3: Metabotropic Receptors

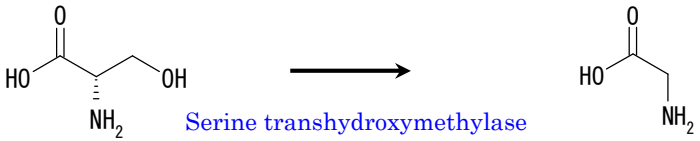
Amino Acid

Glutamic Acid



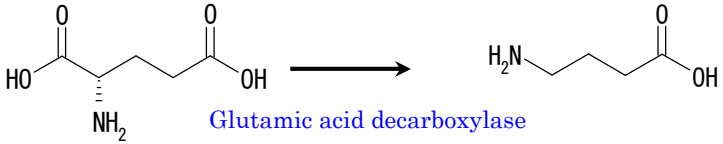
L-Glutamine — L-Glutamic Acid

Glycine



L-Serine — Glycine

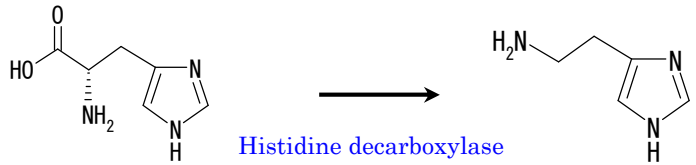
GABA



L-Glutamic Acid — GABA

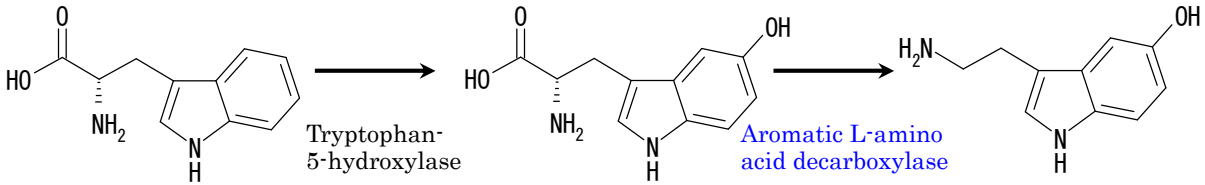
Amine

Histamine



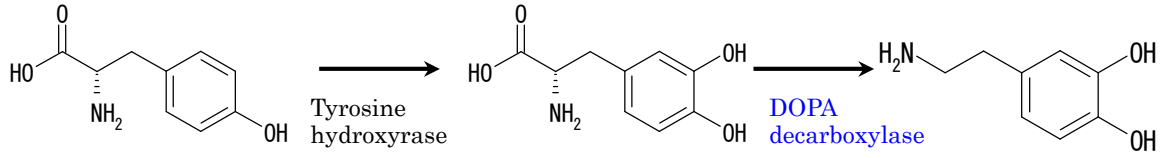
L-Histidine — Histamine

5-HT (Serotonin)

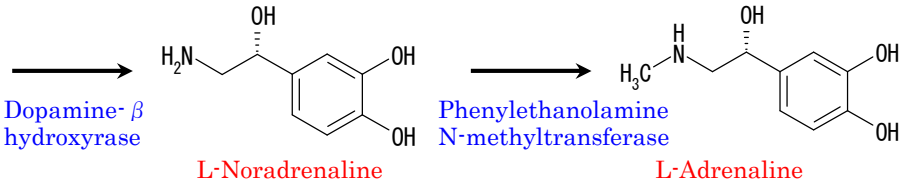


L-Tryptophan — 5-Hydroxy-L-tryptophan — 5-HT (Serotonin)

Dopamine, Noradrenaline, Adrenaline



L-Tyrosine — L-DOPA — Dopamine

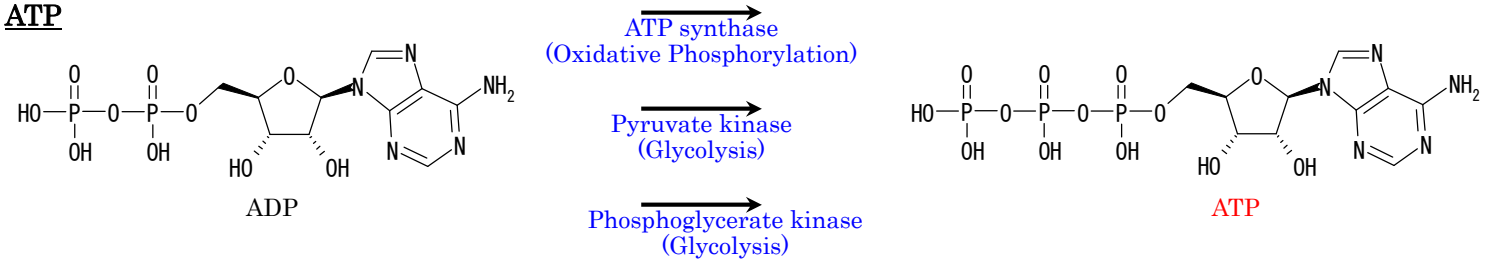


L-Noradrenaline — L-Adrenaline

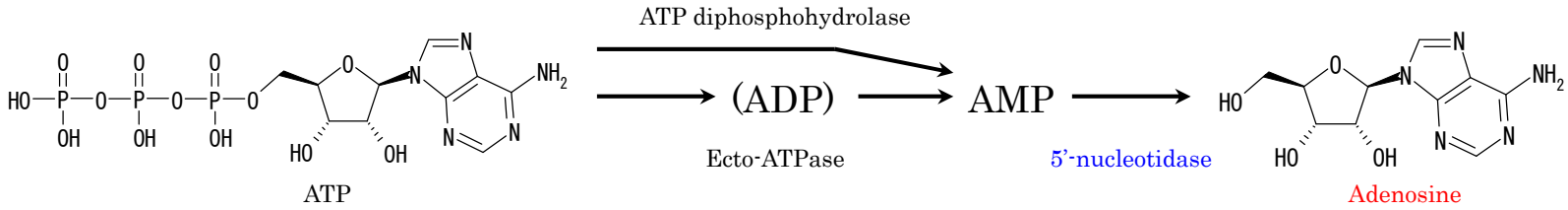
Fig. 3.4: Neurotransmitters and their metabolic pathways. The names of neurotransmitters and synthases are colored red and blue, respectively.

Purine

ATP

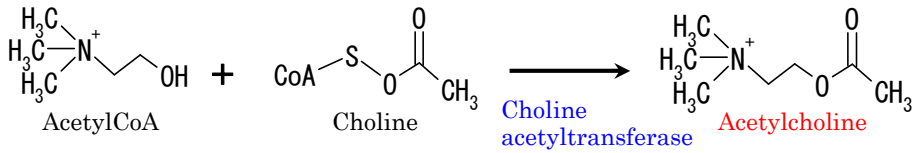


Adenosine

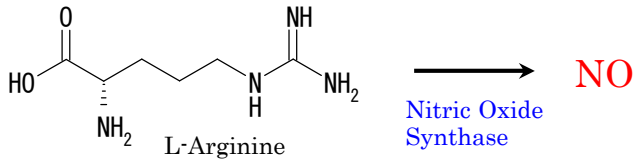


Others

Acetylcholine



NO



Taurine

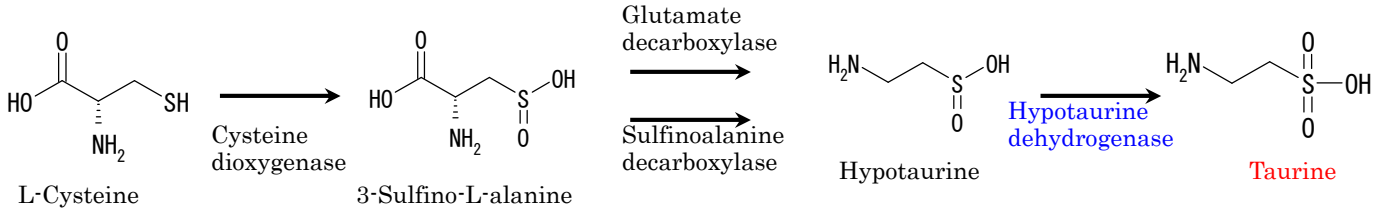


Fig. 3.4 (continued)

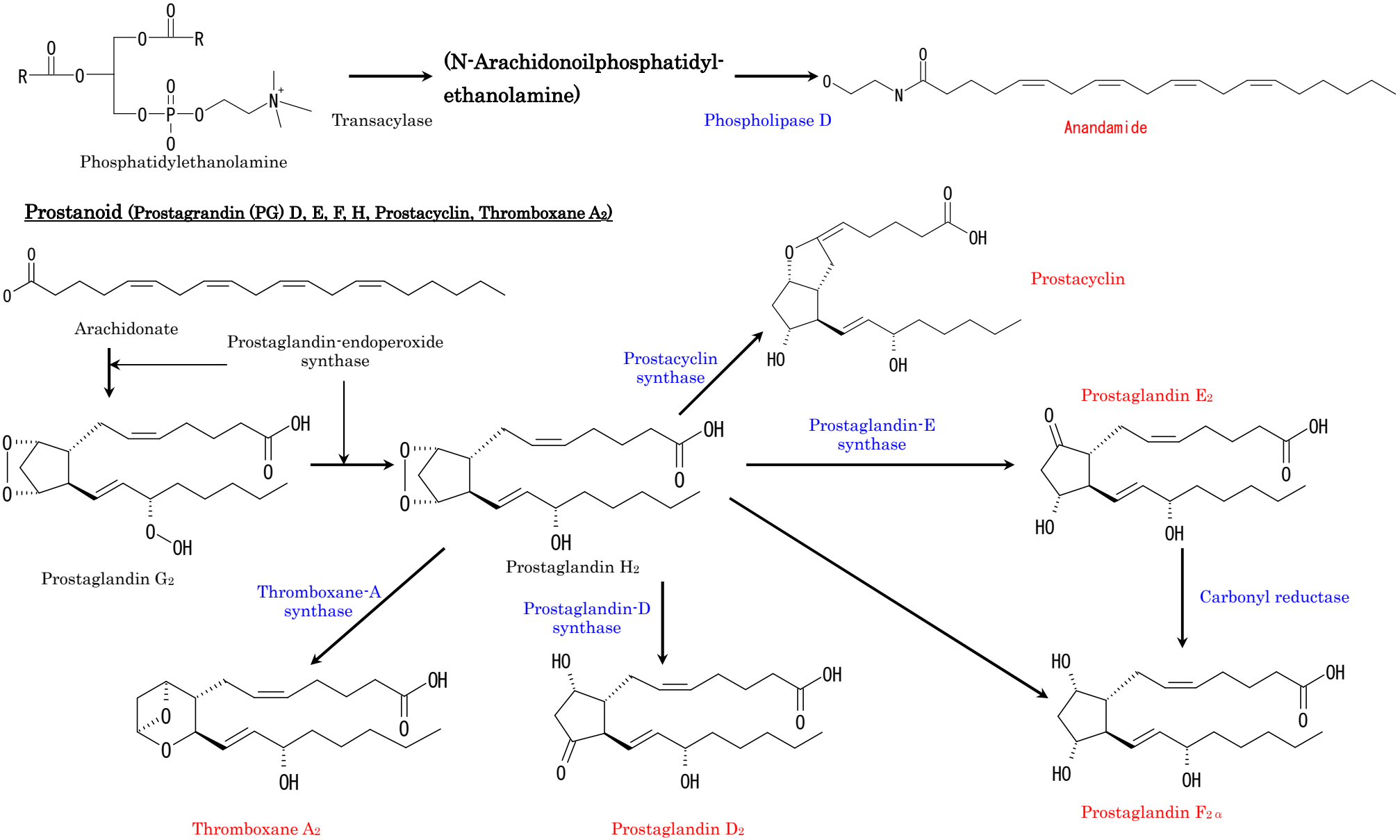


Fig. 3.4 (continued)

3.3.2 Phylogenetic Analysis of Receptors

I then analyzed the relationships between amino acid sequences and functions of receptors. Domain compositions of receptors which I analyzed are shown in Fig. 3.5. Each domain is referred from Pfam database (blue colored names in Fig. 3.5). Some domains are shared among some proteins. For example, ANF_receptor domain is shared among 1.2 receptor, 2.3 receptor, Leu/Ile/Val-binding protein (bacterial amino acid binding protein), and membrane-bound guanylyl cyclase (hormone receptor). ANF_receptor domain is a ligand binding domain in 2.3 receptor, bacterial amino acid binding protein and membrane-bound guanylyl cyclase (Kunishima et al. 2000; McNicoll et al. 1996; Sacks et al. 1989a; van den Akker et al. 2000). ANF_receptor domain in 1.2 receptor is not the ligand binding domain, but the subunit association domain (Ayalon and Stern-Bach 2001). The ligand binding domains of 1.2 receptor are S1 and S2 domains which are defined by Pfam (Armstrong and Gouaux 2000; Armstrong et al. 1998). Glutamine-binding protein, another bacterial amino acid binding protein, also has S1 and S2 ligand binding domains (Hsiao et al. 1996; Sun et al. 1998), and these domains in the bacterial amino acid binding protein are defined as SBP_bac_3 in Pfam database. Soluble guanylyl cyclase and membrane-bound guanylyl cyclase share Guanylate_cyc domain, a catalytic domain.

1.1, 1.4 and 2.1 receptors do not share domains with other receptors I analyzed. Ligand binding sites of 1.1, 1.4 and 2.1 are in Neur_chan_LBD, P2X_receptor and 7tm_1, and transmembrane domains are in Neur_chan_memb, P2X_receptor and 7tm_1, respectively.

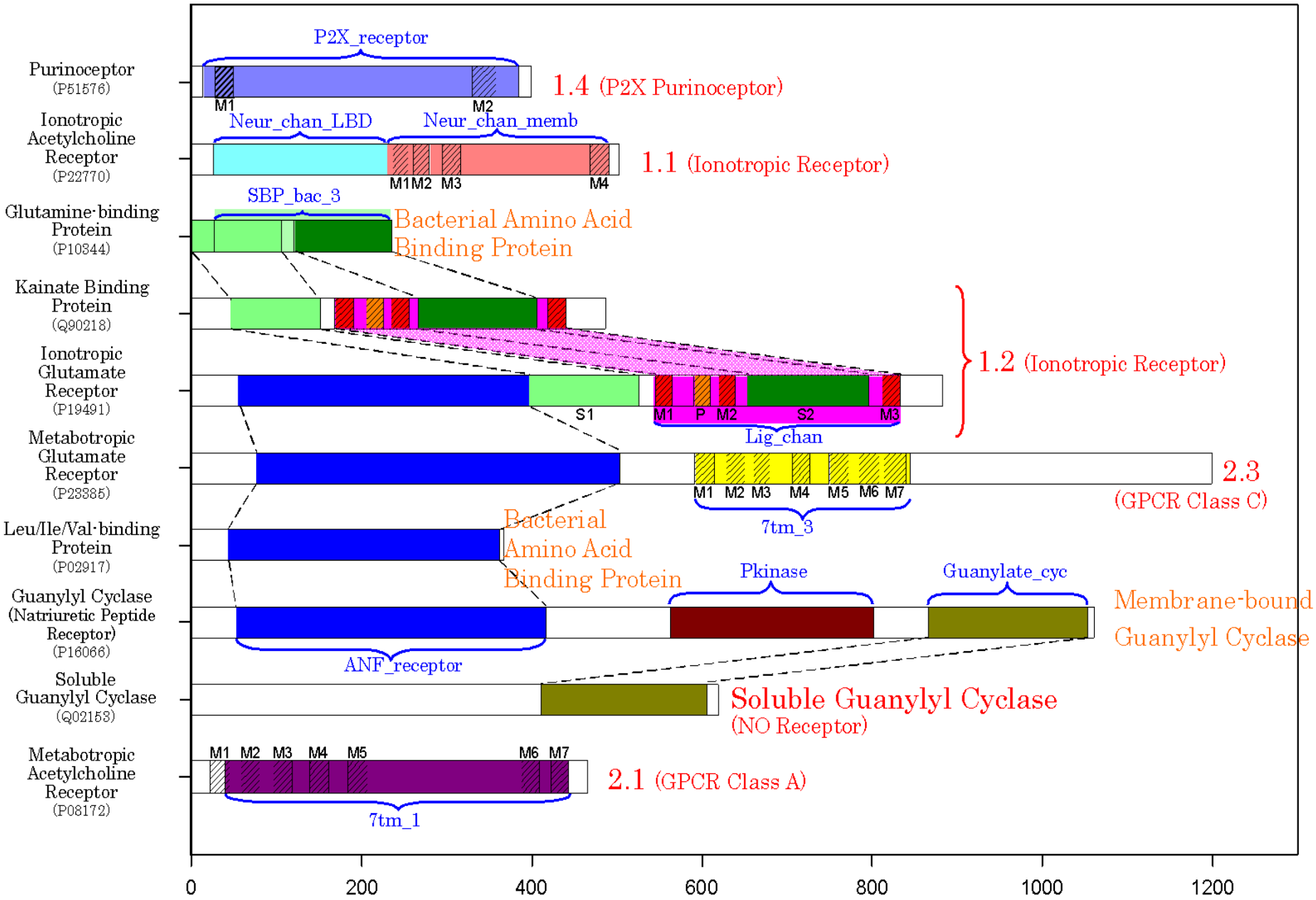


Fig. 3.5: Domain compositions of neurotransmission receptors (red colored names) and their homologous proteins (orange colored names). Domains defined by Pfam are written by blue characters.

3.3.2.1 Phylogenetic Analysis of 1.2 and 2.3 Receptors and Soluble Guanylyl Cyclase

3.3.2.1.1 Domain Changes of 1.2 and 2.3 Receptors and Soluble Guanylyl Cyclase

As I described in the last section, ANF_receptor domain is shared among 1.2 receptor, 2.3 receptor, bacterial amino acid binding protein and membrane-bound guanylyl cyclase (Fig. 3.5). And one of membrane-bound guanylyl cyclase domains, Guanylate_cyc domain, is shared between membrane-bound guanylyl cyclase and soluble guanylyl cyclase which is nitric oxide receptor (Fig. 3.5). These domain compositions suggest that the three receptors are evolutionarily related. I constructed molecular phylogenetic trees by using each domain to pursue the domain changes of 1.2 and 2.3 receptors and soluble guanylyl cyclase through evolution.

3.3.2.1.2 Domain Changes of 1.2 and 2.3 Receptors –Phylogenetic Analysis of ANF_receptor Domain-

Fig. 3.6 shows the phylogenetic tree of 1.2 receptor, 2.3 receptor, bacterial amino acid binding protein and membrane-bound guanylyl cyclase constructed by using ANF_receptor. Bacterial amino acid binding protein exists in bacteria, and the other proteins exist in eukaryotes. Therefore, I defined bacteria amino acid binding protein as the out-group. If I think parsimoniously, the hypothetical ancestral protein was the protein which might have only ANF_receptor domain (Fig. 3.6). I assume that the ancestral protein was the soluble ligand binding protein like bacterial amino acid binding protein. My hypothesis of the evolution from the ancestral protein is as follows (Fig. 3.6). The ionotropic 1.2 receptor was produced by the insertions of the red colored three transmembrane domains (M1-M3 regions in Fig. 3.5), the orange colored

intramembrane domain (P region in Fig. 3.5), the light green and green colored ligand binding domains. Although the ligand binding domains in membrane-bound guanylyl cyclase, 2.3 receptor and bacterial amino acid binding protein are ANF_receptor domains, the ligand binding domain in 1.2 receptor is not ANF_receptor. The ligand binding domains of 1.2 receptor are the light green colored S1 domain and the green colored S2 domains (Fig. 3.5). The ligand binding function of ANF_receptor in 1.2 receptor was lost on the lineage to 1.2 receptor. 7tm_3 domain including seven transmembrane regions was inserted on another lineage, and metabotropic 2.3 receptor (GPCR class C) was generated by the insertion. The insertions of Guanylate_cyc and Pkinase domains on the lineage to membrane-bound guanylyl cyclase generated membrane-bound guanylyl cyclase which is hormone receptor.

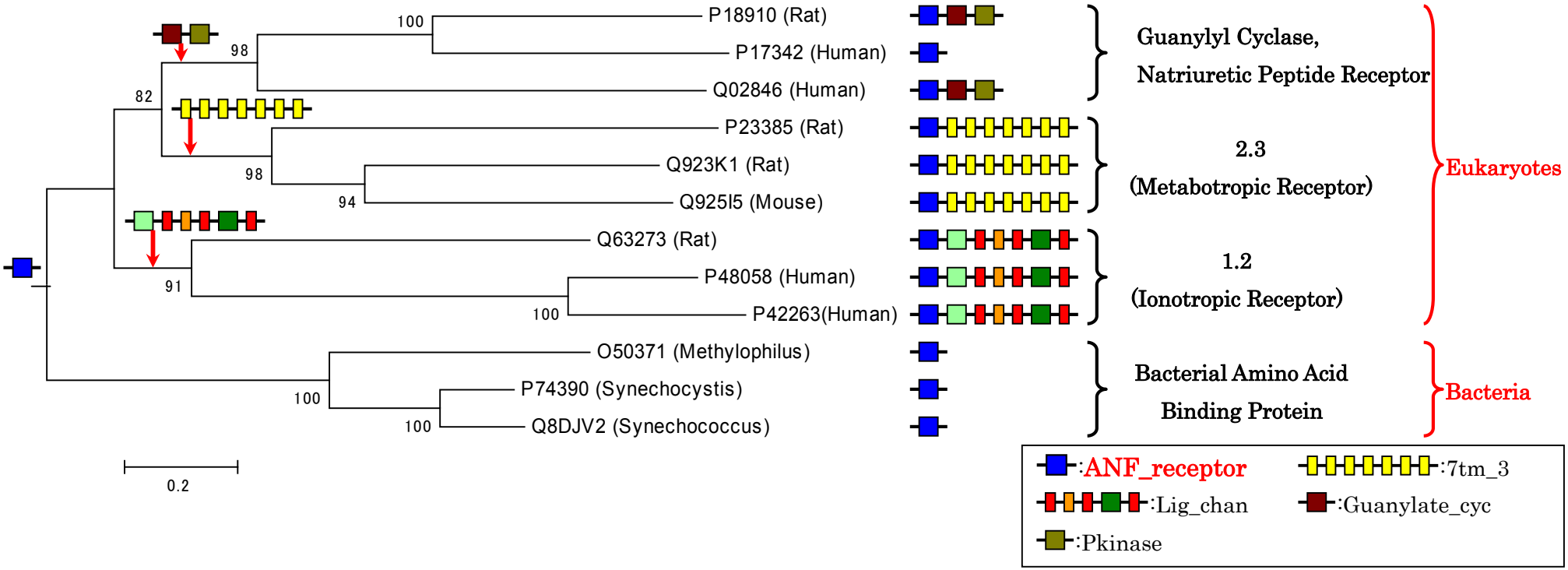


Fig. 3.6: Phylogenetic trees of 1.2 and 2.3 receptors, membrane-bound guanylyl cyclase and bacterial amino acid binding protein constructed by ANF_receptor domains.

3.3.2.1.3 Domain Changes of 2.3 Receptors –Phylogenetic Analysis of 7tm_3 Domain-

I constructed the molecular phylogenetic tree by using 7tm_3 domain which includes seven transmembrane regions and is conserved among all 2.3 receptors (Fig. 3.7). Metabotropic glutamate GABA-like receptor (O96954) exists in porifera, and other receptors exist in eumetazoa (also refer to species distribution in Fig. 3.13). Therefore, I defined metabotropic glutamate GABA-like receptor as the out-group. If I think parsimoniously as I did in the phylogenetic tree of ANF_receptor domain, there are some candidates about the insertions and deletions of ANF_receptors. I postulated that the occurance of the insertion event might be more difficult than that of the deletion. Based on the assumption, I chose the one candidate. It is that the insertion of ANF_receptor may have occurred once, and two deletions of ANF_receptor may have occurred (Fig. 3.7).

If I accept the hypothesis, the evolution of 2.3 receptor family is as follows (Fig. 3.7). The ancestral protein had seven transmembrane regions like metabotropic glutamate GABA-like receptor (O96954). After divergence between porifera and eumetazoa, ANF_receptor domain which is the ligand binding domain was inserted, and ligand binding function was acquired. Then, the ancestor of 2.3 receptor was generated. After that, ANF_receptor domain was lost on the lineage to hypothetical proteins and bride of sevenless protein. Interestingly, bride of sevenless protein is membrane-bound ligand for sevenless tyrosine-kinase receptor (Kramer et al. 1991), that is, 2.3 receptor might change into membrane-bound ligand on the lineage to bride of sevenless protein.

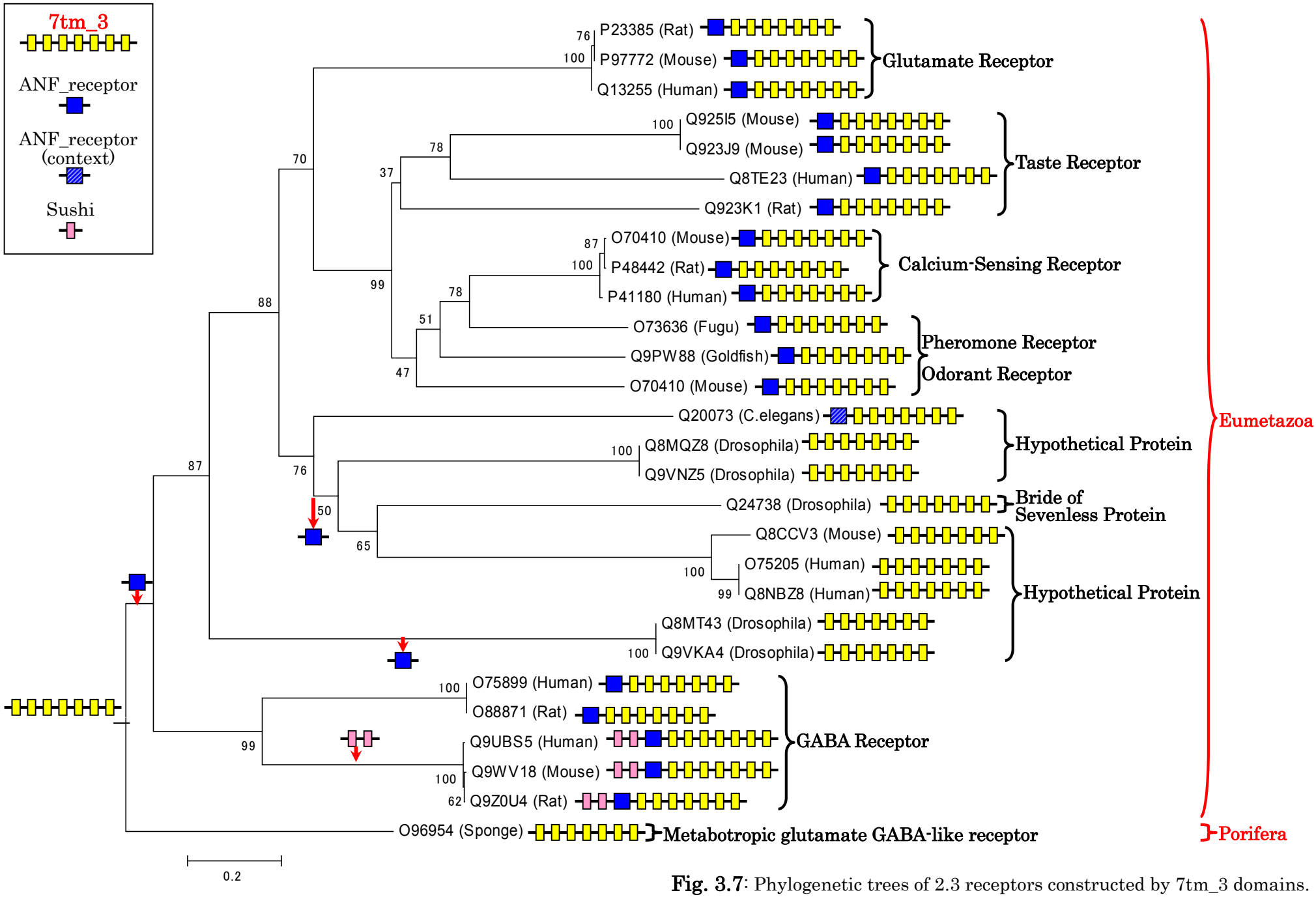


Fig. 3.7: Phylogenetic trees of 2.3 receptors constructed by 7tm_3 domains.

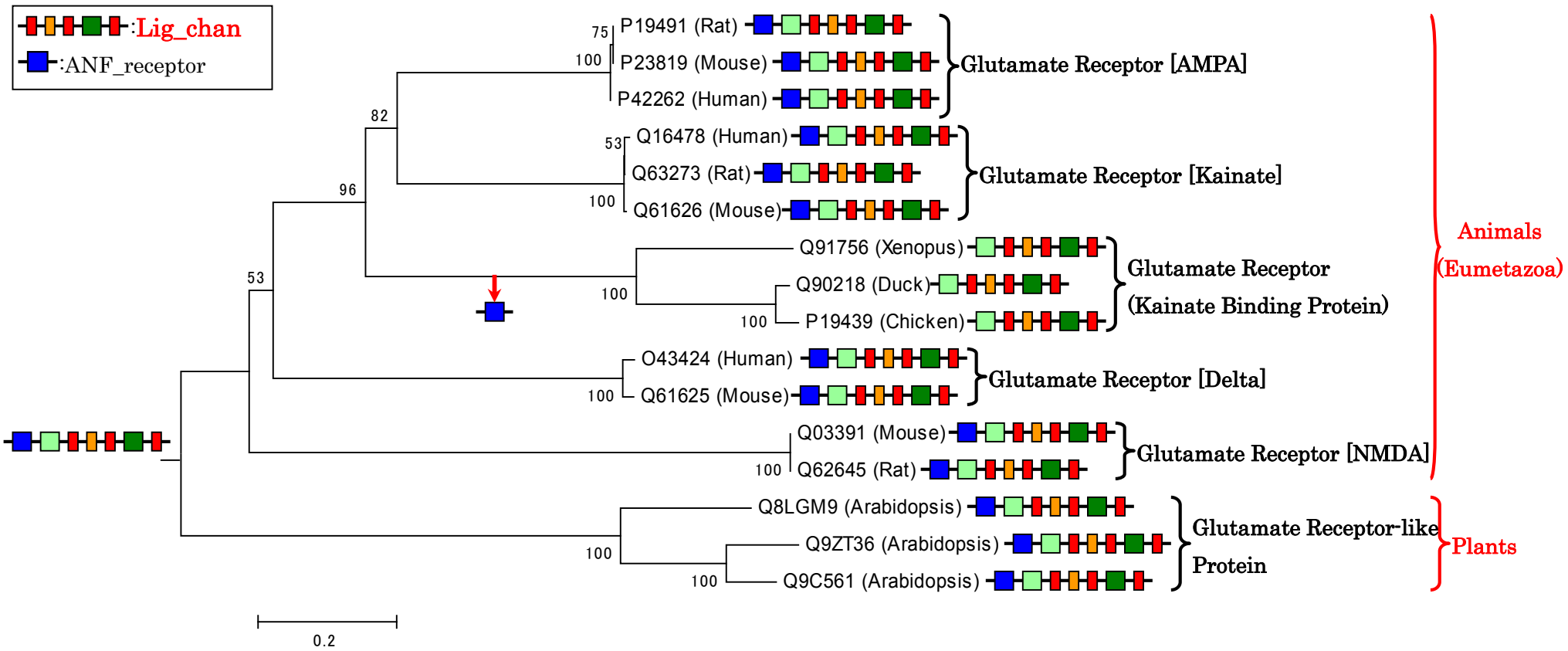


Fig. 3.8: Phylogenetic trees of 1.2 receptors constructed by 7tm_3 domains.

3.3.2.1.4 Domain Changes of 1.2 Receptors –Phylogenetic Analysis of Lig_chan Domain-

Fig. 3.8 is the phylogenetic tree constructed by using Lig_chan domain including three transmembrane domains (M1-M3), one intramembrane domain (P) and one ligand binding domain (S2), but Lib_chan domain does not include the other ligand binding domain (S1) (Fig. 3.5). Glutamate receptor-like proteins exist in plants, and the remaining proteins exist in animals (also refer to species distribution in Fig. 3.14). Therefore, I defined glutamate receptor-like proteins as the out-group. My hypothesis is that the ancestral protein of 1.2 receptors had ANF_receptor domain, Lig_chan domain and S1 domain (Fig. 3.8). And ANF_receptor may have been lost on the lineage to kainate binding protein (Fig. 3.8).

3.3.2.1.5 Domain Changes of Membrane-bound Guanylyl Cyclase –Phylogenetic Analysis of ANF_receptor Domain-

As I described in previous sections, membrane-bound guanylyl cyclase have ANF_receptor domain which also exist in 1.2 and 2.3 receptors. I constructed the phylogenetic tree of membrane-bound guanylyl cyclase by using ANF_receptor domain to know the domain changes of membrane-bound guanylyl cyclase (Fig. 3.9), although this protein is not the receptor I focused on (Table 3.1). I defined bacterial amino acid binding proteins as the out-group, because membrane-bound guanylyl cyclases are present in animals. This tree suggests that Pkinase domain (kinase homology region) and Guanylate_cyc domain (guanylyl cyclase domain) are inserted after the divergence between bacteria and animals. And these domains may have been lost on the lineage to natriuretic peptide receptor C. As a result, natriuretic peptide receptor C might lose the function of guanylyl cyclase.

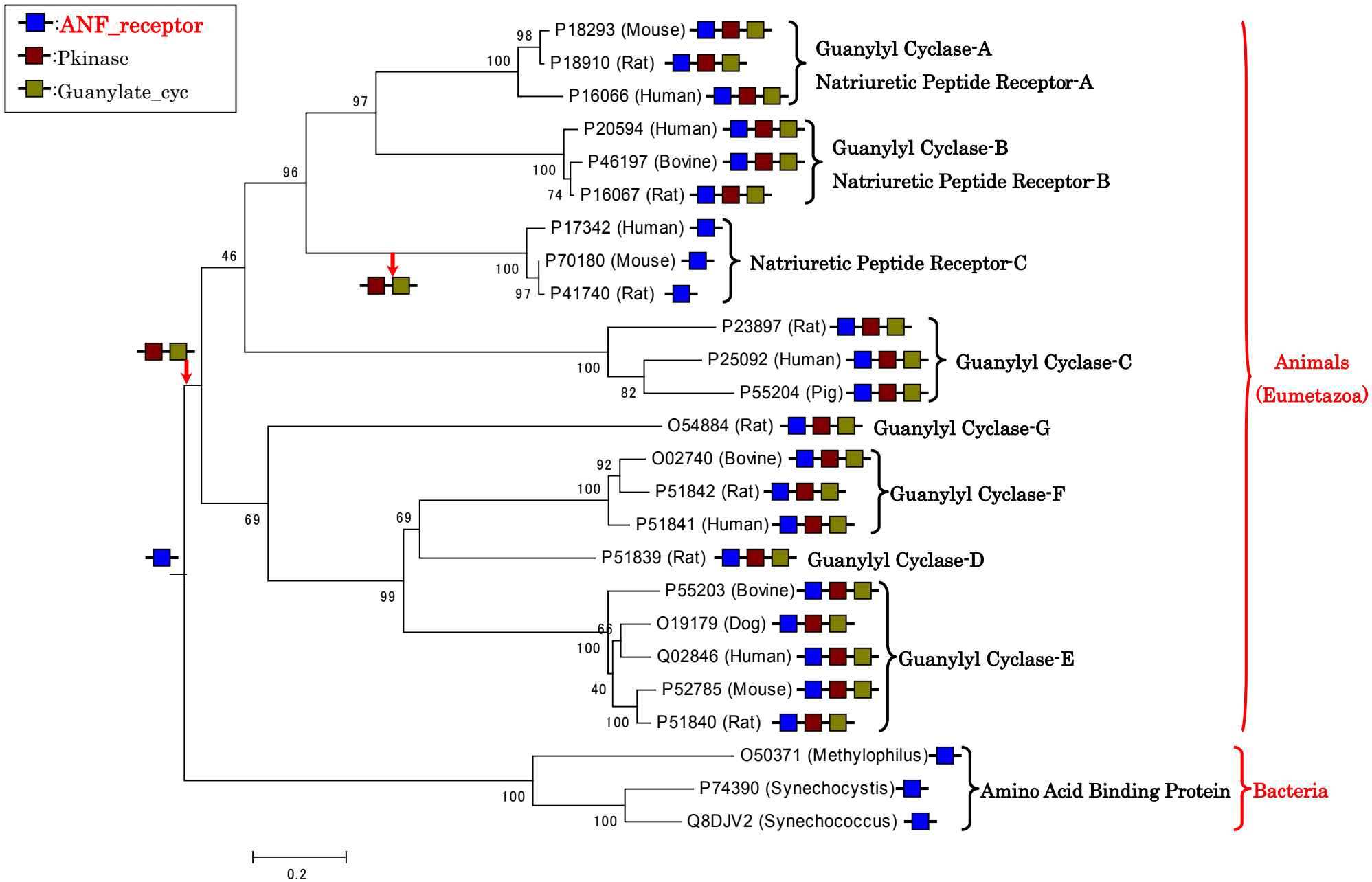


Fig. 3.9: Phylogenetic trees of membrane-bound guanylyl cyclase constructed by ANF_receptor domains.

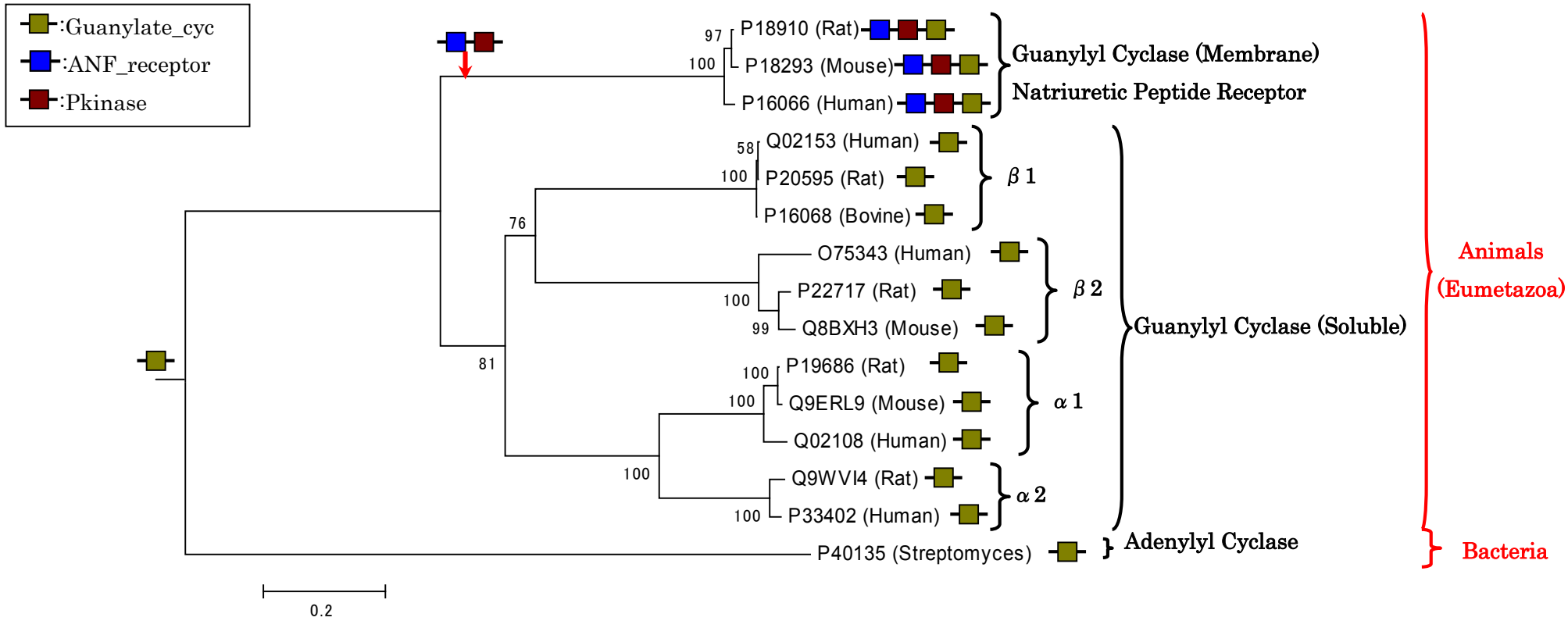


Fig. 3.10: Phylogenetic trees of soluble guanylyl cyclase, membrane-bound guanylyl cyclase and adenylyl cyclase constructed by Guanylate_cyc domains.

3.3.2.1.6 Domain Changes of Soluble Guanylyl Cyclase –Phylogenetic Analysis of Guanylate_cyc Domain-

The phylogenetic tree in Fig. 3.10 was constructed by using all regions of soluble guanylyl cyclase, membrane-bound guanylyl cyclase and adenylyl cyclase which have Guanylate_cyc domain. I defined adenylyl cyclase existed in bacteria as the out-group of this tree. Soluble guanylyl cyclase is the receptor for nitric oxide (Table 3.1). The phylogenetic tree in Fig. 3.10 suggests ANF_receptor and Pkinase domains were inserted after the divergence between bacteria and animals. And after the insertion, soluble guanylyl cyclase may have changed to membrane-bound guanylyl cyclase.

3.3.2.1.7 Domain Changes of 1.2 Receptors –Phylogenetic Analysis of SBP_bac_3, and S1 and S2 Domains-

Fig. 3.11 is the molecular phylogenetic tree constructed by multiple alignment of SBP_bac_3 domain in bacterial amino acid binding protein, and S1 and S2 domains in 1.2 receptors (Fig. 3.5). I defined bacterial amino acid as the out-group, because other proteins exist in plants and animals (also refer to species distribution in Fig. 3.14). This tree suggests that ANF_receptor domain, three red colored transmembrane domains (M1-M3 regions in Fig. 3.5) and one orange colored intramembrane domain (P region in Fig. 3.5) may have been inserted after the divergence between bacteria, and plants and animals. As a result of the insertions, the 1.2 receptors may have been generated.

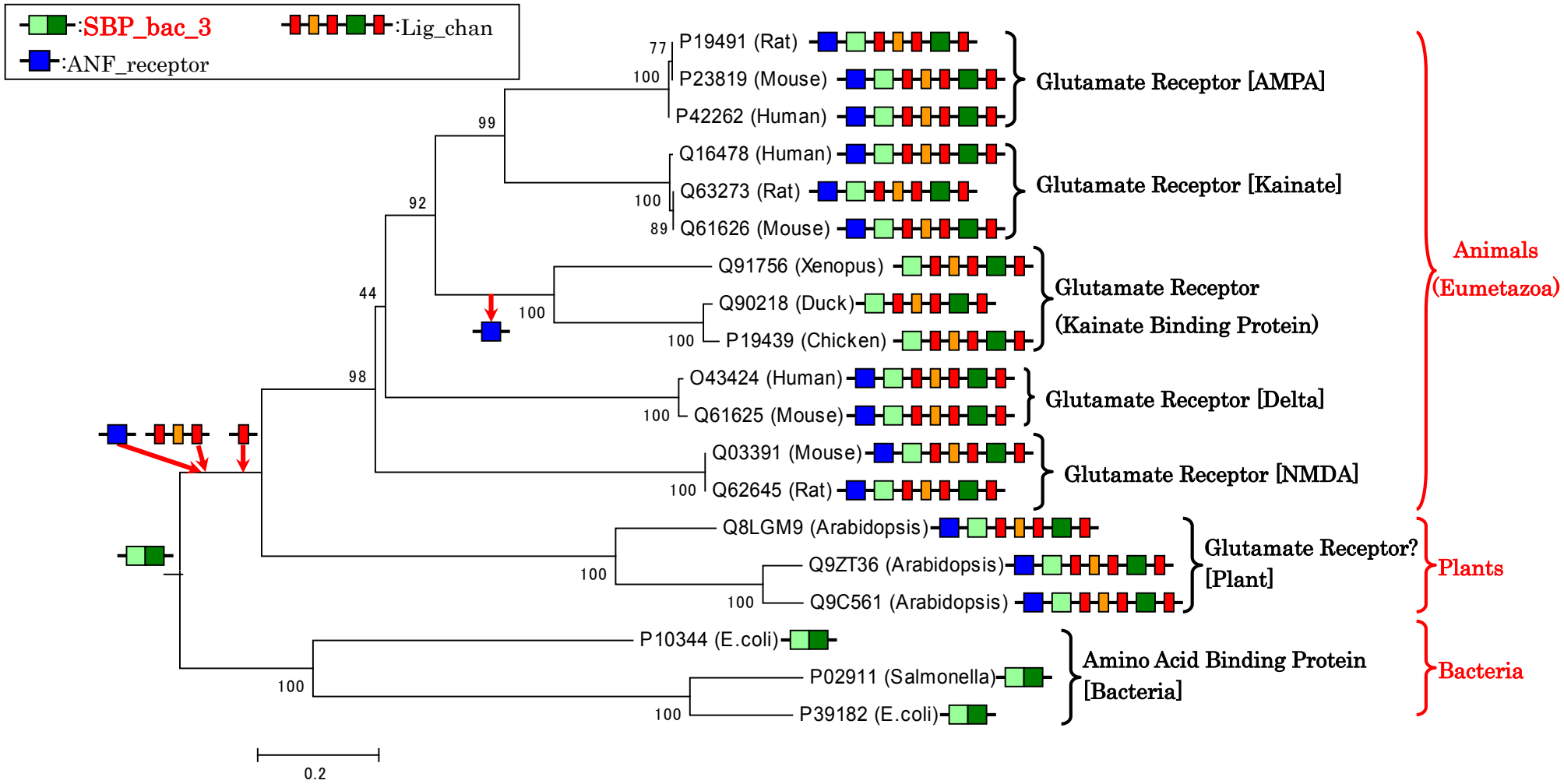


Fig. 3.11: Phylogenetic trees of 1.2 receptors and bacterial amino acid binding protein constructed by S1 and S2 regions in 1.2 receptors (Fig. 3.2) and SBP_bac_3 domain in bacterial amino acid binding protein.

3.3.2.1.8 Domain Changes of 1.2 and 2.3 Receptors and Soluble Guanylyl Cyclase - Construction of Composite Gene Tree -

I hypothesized domain changes of 1.2 and 2.3 receptors and soluble guanylyl cyclase by the molecular phylogenetic trees (Fig. 3.6 – 3.11). I composed the trees (Fig. 3.6 – 3.11) to infer the overall domain change of 1.2 and 2.3 receptors and soluble guanylyl cyclase (Fig. 3.12). The ancestral proteins of the three neurotransmitter receptors may have been generated before the divergence between bacteria and eukaryotes, much before the emergence of the nervous system. And the ancestral proteins might combine some domains, and evolved to 1.2 or 2.3 receptors.

My hypothesis about the evolution of 1.2 receptor is as follows. The two ligand binding domains (light green and green domains in Fig. 3.12) were inserted before the divergence between plants and animals. The ligand binding domains were separated when they were inserted. ANF_receptor domain was also inserted before the divergence between plants and animals. Although ANF_receptor domains in 2.3 receptor, membrane-bound guanylyl cyclase and bacterial amino acid have ligand binding function, ANF_receptor domain in 1.2 receptor lost the function.

I hypothesized that 2.3 receptor had been generated by the combination of ANF_receptor domain including ligand binding sites and 7tm_3 domain including seven transmembrane regions after the divergence between porifera and eumetazoa.

Soluble guanylyl cyclase, receptor for nitric oxide, seems to have been generated without domain changes. A smaller number of amino acid substitutions might produce soluble guanylyl cyclase after the divergence between porifera and eumetazoa.

Although some domains in 1.2 and 2.3 receptors and soluble guanylyl cyclase

may have been generated before the divergence between bacteria and eukaryotes, these receptors might diversify after the divergence between porifera and eumetazoa by a small number of amino acid substitutions.

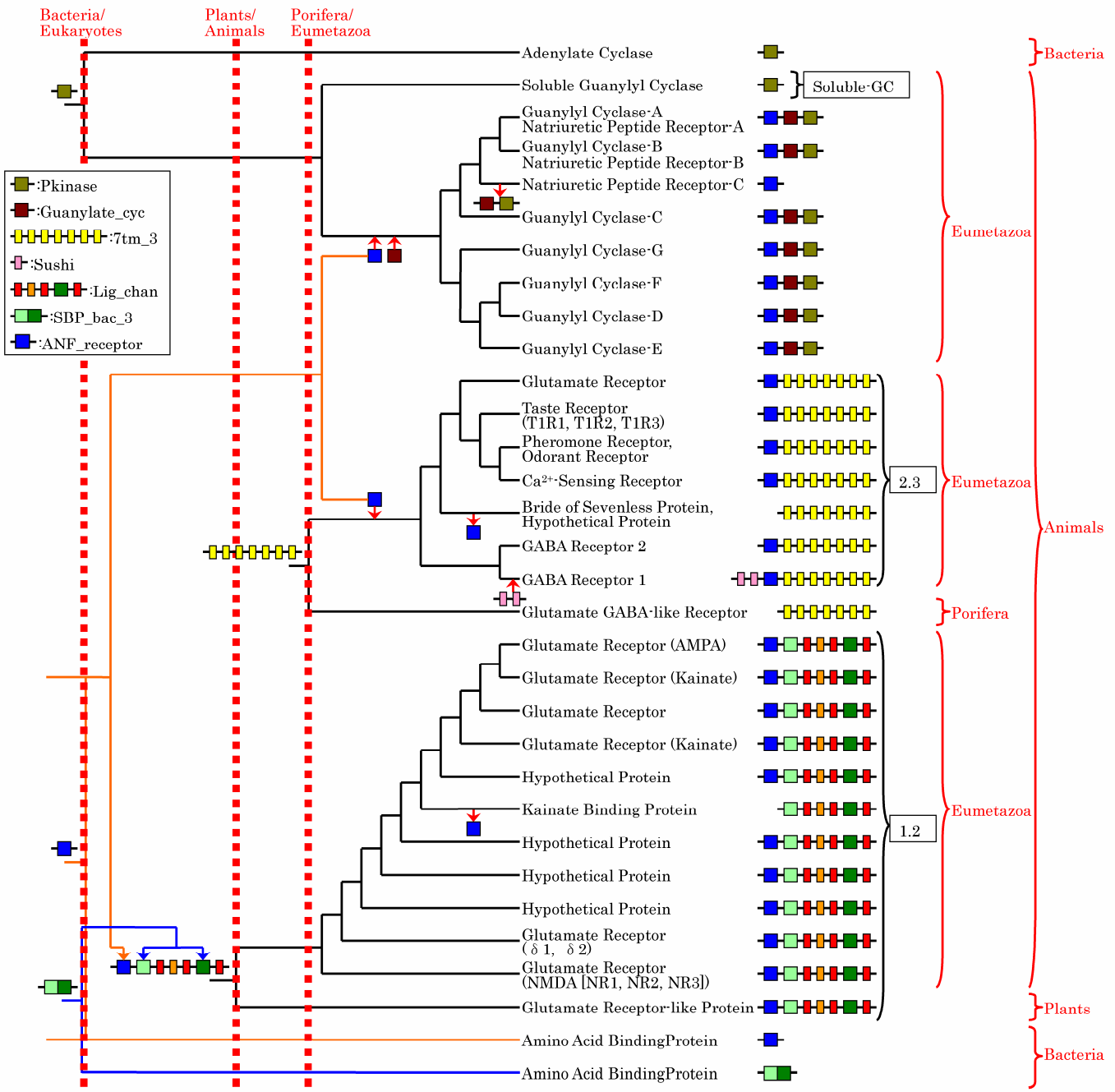
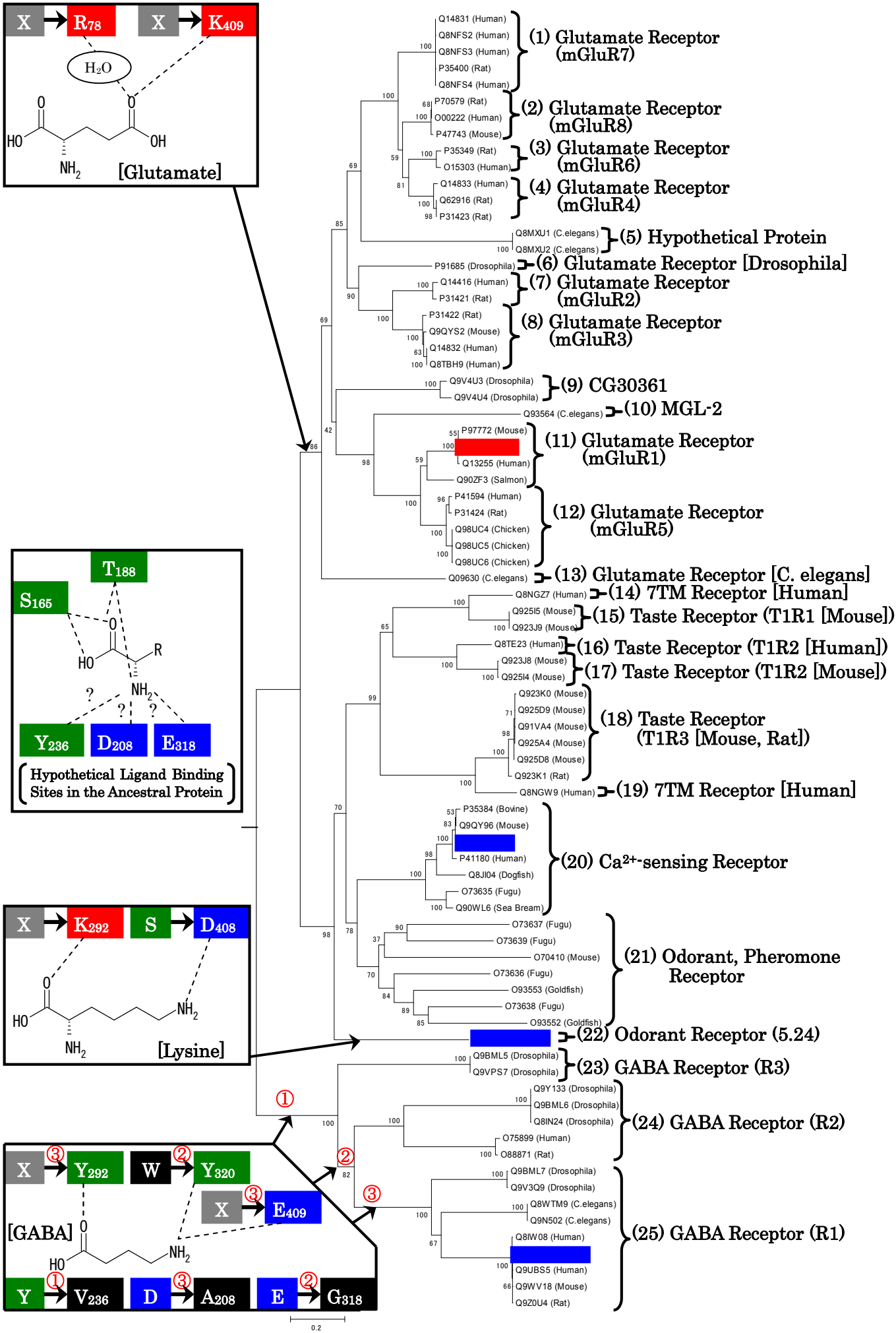


Fig. 3.12: Composite tree of Figs. 3.3-3.8. Branch length has no meaning.

3.3.2.1.9 Necessary Amino Acid Changes for Ligand Binding Specificity of 2.3 Receptors

2.3 receptors are categorized into GPCR class C. 2.3 receptors are dimer proteins, and each subunit binds ligand. But one subunit of GABA receptor which is heterodimer protein may not bind ligand (Kniazeff et al. 2002). Fig. 3.13 shows the phylogenetic tree of 2.3 receptors constructed by using ANF_receptor domains which are ligand binding domains. The important amino acid sites for ligand binding are listed and aligned in Table 3.2. These sites are detected by tertiary structure studies, computer modeling studies and mutation studies (bibliographies cited are listed in Table 3.2). The meaning of color in each amino acid site in Fig. 3.13 is same as the meaning written in Table 3.2. Amino acid numbers I'll describe in this section correspond to the amino acid numbers of metabotropic glutamate receptor 1 of rat (P23385) (Table 3.2).

Fig. 3.13 (next page): Phylogenetic tree of 2.3 receptors constructed by ANF_receptor domains. Protein whose ligand binding sites are detected in the tertiary structure is colored red, proteins whose ligand binding sites are predicted by modeling and experiments are colored blue, respectively. Amino acid substitutions important for changing ligand binding specificities are shown (corresponding sites are surrounded by yellow lines in Table 3.2).


X R78 X K409
H2O
HO OH
NH2
[Glutamate]
Q14831 (Human)
Q8NFS2 (Human)
Q8NFS3 (Human)
P35400 (Rat)
Q8NFS4 (Human)
(1) Glutamate Receptor (mGluR7)
P70579 (Rat)
O00222 (Human)
P47743 (Mouse)
(2) Glutamate Receptor (mGluR8)
P35349 (Rat)
O15303 (Human)
(3) Glutamate Receptor (mGluR6)
Q14833 (Human)
Q62916 (Rat)
P31423 (Rat)
(4) Glutamate Receptor (mGluR4)
Q8MXU1 (C.elegans)
Q8MXU2 (C.elegans)
(5) Hypothetical Protein
P91685 (Drosophila)
(6) Glutamate Receptor [Drosophila]
Q14416 (Human)
P31421 (Rat)
(7) Glutamate Receptor (mGluR2)
P31422 (Rat)
Q9QYS2 (Mouse)
Q14832 (Human)
Q8TBH9 (Human)
(8) Glutamate Receptor (mGluR3)
Q9V4U3 (Drosophila)
Q9V4U4 (Drosophila)
(9) CG30361
Q93564 (C.elegans)
(10) MGL-2
P97772 (Mouse)
Q13255 (Human)
Q90ZF3 (Salmon)
(11) Glutamate Receptor (mGluR1)
P41594 (Human)
P31424 (Rat)
Q98UC4 (Chicken)
Q98UC5 (Chicken)
Q98UC6 (Chicken)
(12) Glutamate Receptor (mGluR5)
Q09630 (C.elegans)
(13) Glutamate Receptor [C. elegans]
T188
S165
HO R
NH2
? ? ?
Y236 D208 E318
Hypothetical Ligand Binding Sites in the Ancestral Protein
Q8NGZ7 (Human)
(14) 7TM Receptor [Human]
Q925I5 (Mouse)
Q923J9 (Mouse)
(15) Taste Receptor (T1R1 [Mouse])
Q8TE23 (Human)
(16) Taste Receptor (T1R2 [Human])
Q923J8 (Mouse)
Q925I4 (Mouse)
(17) Taste Receptor (T1R2 [Mouse])
Q923K0 (Mouse)
Q925D9 (Mouse)
Q91VA4 (Mouse)
Q925A4 (Mouse)
Q925D8 (Mouse)
Q923K1 (Rat)
(18) Taste Receptor (T1R3 [Mouse, Rat])
Q8NGW9 (Human)
(19) 7TM Receptor [Human]
P35384 (Bovine)
Q9QY96 (Mouse)
P41180 (Human)
Q8JI04 (Dogfish)
O73635 (Fugu)
Q90WL6 (Sea Bream)
(20) Ca2+-sensing Receptor
X K292 S D408
HO NH2
NH2
[Lysine]
O73637 (Fugu)
O73639 (Fugu)
O70410 (Mouse)
O73636 (Fugu)
O93553 (Goldfish)
O73638 (Fugu)
O93552 (Goldfish)
(21) Odorant, Pheromone Receptor
(22) Odorant Receptor (5.24)
Q9BML5 (Drosophila)
Q9VPS7 (Drosophila)
(23) GABA Receptor (R3)
Q9Y133 (Drosophila)
Q9BML6 (Drosophila)
Q8IN24 (Drosophila)
O75899 (Human)
O88871 (Rat)
(24) GABA Receptor (R2)
Q9BML7 (Drosophila)
Q9V3Q9 (Drosophila)
Q8WTM9 (C.elegans)
Q9N502 (C.elegans)
Q8IW08 (Human)
Q9UBS5 (Human)
Q9WV18 (Mouse)
Q9Z0U4 (Rat)
(25) GABA Receptor (R1)
① ② ③
X Y292 W Y320
X E409
[GABA]
HO NH2
Y V236 D A208 E G318
0.2

Amino Acid Site Number of mGluR1 (P23385)		74	78	164	165	186	188	208	236	292	293	318	320	323	408	409	References
1	Glutamate Receptor (mGluR7)	N	R	G	S	A	T	D	Y	N	D	D	W	K	G	K	
2	Glutamate Receptor (mGluR8)	K	R	A	S	A	T	D	Y	N	E	D	W	K	G	K	
3	Glutamate Receptor (mGluR6)	Q	R	A	S	A	T	D	Y	N	E	D	W	K	G	K	
4	Glutamate Receptor (mGluR4)	K	R	G	S	A	T	D	Y	N	E	D	W	K	G	K	
5	Hypothetical Protein	Q	R	G	S	S	T	D	Y	D	E	D	W	K	S	Y	
6	Glutamate Receptor (Drosophila)	R	R	Y	S	A	T	D	Y	R	A	D	W	Q	S	K	
7	Glutamate Receptor (mGluR2)	R	R	Y	S	A	T	D	Y	R	S	D	W	L	S	K	
8	Glutamate Receptor (mGluR3)	R	R	Y	S	A	T	D	Y	R	S	D	W	Q	S	K	
9	CG30361	G	A	S	S	F	T	D	Y	S	D	D	W	R	D	Q	
10	MGL-2	Y	R	K	S	S	T	D	Y	E	G	D	W	R	P	K	
11	Glutamate Receptor (mGluR1)	Y**	R*	S*	S**	S**	T**	D**	Y**	E*	G*	D**	W	R**	S	K**	Kunishima et al. 2000
12	Glutamate Receptor (mGluR5)	Y	R	S	S	S	T	D	Y	E	G	D	W	R	S	K	
13	Glutamate Receptor (C.elegans)	R	R	Y	S	A	T	D	Y	G	T	E	W	N	D	K	
14	7TM Receptor (Human)	H	W	S	T	A	S	D	Y	S	R	E	W	S	S	S	
15	Taste Receptor (T1R1 [Mouse])	H	L	N	T	E	S	D	Y	N	R	E	W	S	S	A	
16	Taste Receptor (T1R2 [Human])	I	L	N	S	S	I	A	Y	P	D	E	W	D	R	V	
17	Taste Receptor (T1R2 [Mouse])	L	L	N	S	S	I	A	Y	P	E	E	W	D	R	V	
18	Taste Receptor (T1R3 [Mouse, Rat])	L	L	S	S	S	S	D	Y	S	A	E	W	S	H	Q	
19	7TM Receptor (Human)	N	L	S	S	G	S	D	Y	S	V	E	W	S	H	Q	
20	Ca^{2+}-sensing Receptor	R	W	G	S+	A	S+	D+	Y+	S	G	E+	W	S	R	I	Zhang et al. 2002
21	Odorant, Pheromone Receptor	H/R/L	F/H/L /W/Y	E/G/ L/T	S/T	F/S/ T/Y	S/T	D	Y	A/H/ S/V	D/E/F/ H/S/Q	E	W	A/D/S /T/-	R/P	I/P/ Y	
22	Odorant Receptor (5.24)	K	Q	S	S+	A	T+	D	Y+	K+	S	D+	W	S	D+	M	Kuang et al. 2003
23	GABA Receptor (R3)	D	E	C	S	G	T	D	V	S	Q	E	M	P	Q	Y	
24	GABA Receptor (R2)	I/T	V	C	T/P	A	T	D/E	R	N/D	E/Q	A/G	Y	D/S/A	K/R	F	
25	GABA Receptor (R1)	G	C	C	S+	G	S	A	V	Y+	V/E	G	Y+	N/T	Q/P	E+	Kniazeff et al. 2002

:Basic Amino Acid (Positive Charged)

:Acidic Amino Acid (Negative Charged)

:Uncharged Polar Amino Acid

:Nonpolar Amino Acid

** Ligand binding site which interacts directly with ligand detected by tertiary structure study.

* Ligand binding site which interacts indirectly with ligand detected by tertiary structure study.

+ Ligand binding site which interacts with ligand predicted by modeling and experiments.

Table 3.2: Important sites for ligand binding in 2.3 receptors

Among 2.3 receptors, amino acid sites of 165, 188, 208, 236, and 318 are conserved (red arrows in Table 3.2). And these sites are also conserved in bacterial amino acid binding proteins which are the out-group proteins of 2.3 receptors (Fig. 3.6). I hypothesized that amino acids of these sites in the ancestral protein of 2.3 receptors are S165, T188, D208, Y236, E318. S165 directly interacts with α-carboxyl group of ligand, and T188 directly interact with α-carboxyl group and α-amino group of ligand in bacterial amino acid binding protein (Sack et al. 1989a; Sack et al. 1989b) and metabotropic glutamate receptor (Kunishima et al. 2000). Therefore, I hypothesized that the ancestral protein had possessed the ability of binding α-carboxyl and α-amino groups of ligand (Fig. 3.13). Although D208, Y236 and E318 in bacterial amino acid binding protein does not interact with ligand (Sack et al. 1989a; Sack et al. 1989b), these corresponding sites in metabotropic glutamate receptor (D208, Y236 and D318) directly interact with α-amino group of glutamate (Kunishima et al. 2000). Therefore, I hypothesized that D208, Y236 and E318 in the ancestral protein had interacted with α-amino group of ligand. The ancestral protein of 2.3 receptor may have already acquired the ability of binding α-carboxyl group and α-amino group of ligand from these hypotheses, that is, the ligand of the ancestral protein might be amino acid.

Tertiary structure analysis detected ligand binding sites of metabotropic glutamate receptor (Kunishima et al. 2000). These sites are shown in Table 3.2 (amino acid sites with asterisks). I predicted the branches where these amino acid sites appeared by amino acid substitutions. R78 and K409 are conserved among all glutamate receptors, and I predicted amino acid substitutions to R78 and K409 had occurred on the branch to glutamate receptors (Fig. 3.13). After these two substitutions, glutamate receptors may have been generated. Therefore, these two amino acid

substitutions are essential substitutions for producing glutamate receptors. The amino group of K409 directly interacts with the γ-carboxyl group of glutamate ligand, and the amino group of R78 indirectly interacts with the γ-carboxyl group of glutamate ligand (Kunishima et al. 2000) (Fig. 3.13). The amino acid substitutions to the two basic amino acids might occur before the divergence between vertebrates and invertebrates and enable 2.3 receptor to bind glutamate (Fig. 3.13).

Odorant 5.24 receptor is the receptor for lysine. Ligand binding sites of odorant receptor 5.24 were predicted by computer modeling and mutation study (Kuang et al. 2003) (Table 3.2). Although odorant receptor 5.24 is not the receptor listed in Table 3.1, I predicted the essential amino acid substitutions for generating the receptor. Predicted ligand binding sites except for K292 and D408 may have conserved from the ancestral protein of 2.3 receptor, if I think parsimoniously. Therefore, I hypothesized amino acid substitutions to K292 which may interact with α-carboxyl group of lysine and D408 which may interact with ε-amino group are important substitutions for generating odorant receptor 5.24. Especially, amino acid substitution to D408 seems to be more important. Amino acids which can interact with α-carboxyl group of amino acid may have been conserved from the ancestral protein of 2.3 receptor by my prediction. Amino acid substitution to D408 enabled 2.3 receptor to interact with ε-amino group of lysine. Therefore, I hypothesized the essential amino acid substitution for generating odorant 5.24 receptor was the amino acid substitution to D408.

There are three kinds of subunits in GABA receptor; R1, R2 and R3. R1 binds GABA, but R2 may not bind GABA (Kniazeff et al. 2002). The function of R3 is not well known, and R3 exists only in Drosophila (Mezler et al. 2001). GABA receptor is a heterodimer protein composed of R1 and R2 (Jones et al. 1998; Kaupmann et al. 1998;

White et al. 1998). Ligand binding sites of R1 subunit predicted by computer modeling study (Kuniazeff et al. 2002) are S165, Y292, Y320 and E409 (Table 3.2). S165 may have conserved from the ancestral protein if I think hypothetically (Fig. 3.13). Amino acid substitutions to Y292, Y320 and E409 may have occurred on the lineage to GABA binding receptor (Fig. 3.13). The amino acid substitutions to Y320 and E409 might enable R1 subunit to interact with amino group of GABA (Fig. 3.13). As I described in previous paragraph, D208, Y236 and E318 in the ancestral protein might interact with α-amino group of ligand (Fig. 3.13). I hypothesized that these D208, Y236 and E318 might have changed to A208, V236 and G318 which are nonpolar amino acids on the lineage to R1 subunit (Fig. 3.13). After these substitutions, GABA receptor R1 subunit might have lost the ability to bind α-amino group which the ancestral protein can interact with. I hypothesized that the amino acid substitutions to Y320 and E409 which had enabled 2.3 receptor to interact with amino group of GABA and the substitutions to A208, V236 and G318 which had made 2.3 receptor to lose the function for binding α-amino group are the essential amino acid substitutions for GABA receptor. These essential substitutions might occur before the divergence between vertebrates and invertebrates (Fig. 3.13).

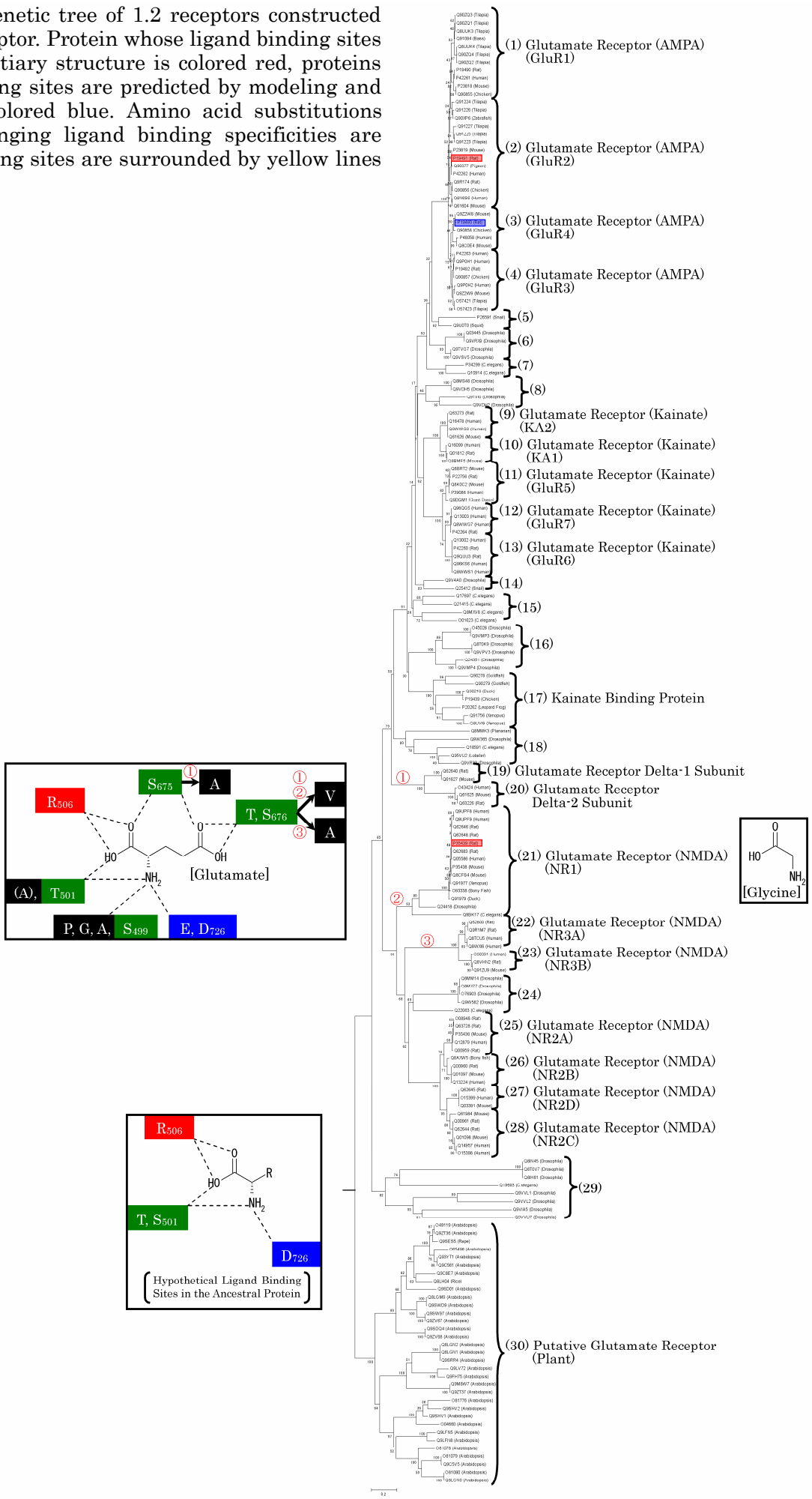
3.3.2.1.10 Necessary Amino Acid Changes for Ligand Binding Specificity of 1.2 Receptors

All of 1.2 receptors are ionotropic glutamate receptors which are tetramer proteins composed of two dimers (Sun et al. 2002), although functions of plants' putative glutamate receptor are not well-known. Delta subunits of glutamate receptor do not bind ligand (Mayat et al. 1995). Heterotetrameric NMDA-type glutamate receptor

composed of NR1 and NR2 requires both glutamate and glycine for efficient activation (Johnson and Ascher 1987; Kleckner and Dingledine 1988; Laube et al. 1998), and that composed of NR1 and NR3 is activated by glycine alone (Chatterton et al. 2002). Ligand for NR2 is glutamate (Laube et al. 1997), and that for NR1 is glycine (Hirai et al. 1996; Kuryatov et al. 1994; Wafford et al. 1995). There are no evidences of NR3 for binding glutamate or glycine, therefore NR3 may not bind ligand. There are different amino acids between ligand binding sites of glutamate binding subunits and corresponding sites of glutamate nonbinding subunits (Table 3.3). And I hypothesized that these amino acid differences produce the different ligand binding functions. Based on the amino acid differences, I hypothesized amino acid substitutions which are essential for the changes of ligand binding functions.

Fig. 3.14 shows the phylogenetic tree of 1.2 receptors constructed by using ligand binding domains. The two divided regions are ligand binding domains in 1.2 receptors (light green and green colored regions of 1.2 receptors in Fig. 3.5). The phylogenetic tree was constructed by the multiple alignment of the two divided ligand binding domains. The important amino acid sites for ligand binding are listed and aligned in Table 3.3. These sites are detected by tertiary structure studies, computer modeling studies and mutation studies (bibliographies cited are listed in Table 3.3). The meaning of color in each amino acid site in Fig. 3.14 is same as the meaning written in Table 3.3. Amino acid numbers I'll describe in this section correspond to the amino acid numbers of AMPA sensitive ionotropic glutamate receptor 2 of rat (P19491) (Table 3.3).

Fig. 3.14: Phylogenetic tree of 1.2 receptors constructed by using ANF_receptor. Protein whose ligand binding sites are detected by tertiary structure is colored red, proteins whose ligand binding sites are predicted by modeling and experiments are colored blue. Amino acid substitutions important for changing ligand binding specificities are shown (corresponding sites are surrounded by yellow lines in Table 3.3).



Amino Acid Site Number of GluR2 (P19491)		499	501	506	671	675	676	724	726	
1	Glutamate Receptor (AMPA) (GluR1)	P	T	R	L	S	T	L	E	
2	Glutamate Receptor (AMPA) (GluR2)	P**	T**	R**	L*	S**	T**	L*	E**	Armstrong and Gouaux 2000
3	Glutamate Receptor (AMPA) (GluR4)	P	T+	R+	L	S	T+	L	E+	Lampinen et al. 1998
4	Glutamate Receptor (AMPA) (GluR3)	P	T	R	L	S	T	L	E	
5	(P26591)	P	T	R	L	S	T	L	E	
6	(Q03445)	A	T	R	L	S	T	L	E	
7	(P34299)	P/S	T	R	Q/N	S	T	L	E	
8	(Q8MS48)	D	T	R	L/K	S/A	T	L	E	
9	Glutamate Receptor (Kainate) (KA2)	A	T	R	I	S	T	L	E	
10	Glutamate Receptor (Kainate) (KA1)	G	T	R	I	S	S	L	E	
11	Glutamate Receptor (Kainate) (GluR5)	P	T	R	V	S	T	L	E	
12	Glutamate Receptor (Kainate) (GluR7)	P	T	R	V	S	T	L	E	
13	Glutamate Receptor (Kainate) (GluR6)	P	A	R	V	S	T	L	E	
14	(Q9V4A0)	P/S	T	R	L/R	S	T	L/F	E	
15	(Q17697)	P/S/A	T	R	L/I	S	T	L/M	E	
16	(O45028)	D	T	R	I/M	S/A	T	L	E	
17	Kainate Binding Protein	P	T	R	L/I	S	T/S	L/I	E	
18	(Q8MMK3)	A/P/S	T/K/S	R	Q/V/I/M/L	S/D/A	I/T/S/A	F/I/L	D/E	
19	Glutamate Receptor Delta-1 Subunit	A	T	R	V	A	V	L	D	
20	Glutamate Receptor Delta-2 Subunit	A	T	R	V	A	V	V	D	
21	Glutamate Receptor (NMDA) (NR1)	P/G**	T**	R**	V	S/N**	V	I	D**	Furukawa and Gouaux 2003
22	Glutamate Receptor (NMDA) (NR3A)	S	S	R	V	S	A	I	D	
23	Glutamate Receptor (NMDA) (NR3B)	S	S	R	V	S	A	I	D	
24	(Q8MM14)	S	M/K	R	V/I	H/N	T	I	D	
25	Glutamate Receptor (NMDA) (NR2A)	S	T	R	V	S	T	I	D	
26	Glutamate Receptor (NMDA) (NR2B)	S	T	R	V	S	T	I	D	
27	Glutamate Receptor (NMDA) (NR2D)	S	T	R	V	S	T	I	D	
28	Glutamate Receptor (NMDA) (NR2C)	S	T	R	V	S	T	I	D	
29	(Q8IN45)	A/G/P	T/I/L/R/F	R/D	R/D/E/Q	F/Y	T/F/N	N/L/V/I/A/Q	D/E/G/Q	
30	Putative Glutamate Receptor (Plant)	-	T/A/S	R	Q/T/P/ F/L/S	F/Y/V/ I/L/M	A/R/V/ T/I	I/V/A/S	D/S/N /M/L	

:Basic Amino Acid (Positive Charged)

:Acidic Amino Acid (Negative Charged)

:Uncharged Polar Amino Acid

:Nonpolar Amino Acid

** Ligand binding site which interacts directly with ligand detected by tertiary structure study.

+ Ligand binding site which interacts with ligand predicted by modeling and experiments.

Table 3.3: Important sites for ligand binding in 1.2 receptors

The three amino acid sites at the positions of 501, 506 and 726 are conserved in both 1.2 receptors and bacterial amino acid binding proteins which are the out-group of 1.2 receptors (Fig. 3.11). These three sites may have been conserved from the ancestral protein of 1.2 receptor and bacterial amino acid binding protein. If I predict parsimoniously based on the amino acid sites 501, 506 and 726 in 1.2 receptors and bacterial amino acid binding protein, the three amino acid sites in the ancestral protein are T or S501, R506, and D726. Based on the tertiary structure of AMPA-sensitive glutamate receptor 2 in rat (Armstrong and Gouaux 2000), T or S501 may interact with α-calboxyl group and α-amino group of ligand (Fig. 3.14). And R506 and D726 may interact with α-calboxyl group and α-amino group of ligand, respectively (Fig. 3.14). Therefore, I hypothesized the ancestral protein have the function of binding amino acid.

I predicted the amino acid substitutions which are essential for changing ligand binding function. As I described in the previous paragraph of this chapter, delta 1 and 2 subunits of glutamate receptor and NR3A-B subunits of NMDA-sensitive glutamate receptor do not bind glutamate. And ligand of NMDA-sensitive glutamate receptor NR1 subunit is glycine. I hypothesized the essential amino acid substitutions for losing glutamate binding function.

The hypothetical essential amino acid substitutions for generating delta subunit 1 and 2 are the amino acid substitutions to A675 and V676. If I think parsimoniously, these substitutions might occur on the lineage to delta 1 and 2 subunits (① in Fig. 3.14). S675 and S676 in rat AMPA sensitive glutamate receptor 2 interact with γ-carboxyl group of glutamate (Armstrong and Gouaux 2000), therefore delta 1 and 2 subunits might lose the binding function of γ-calboxyl group of glutamate by the amino acid substitutions to nonpolar amino acids at the positions of 675 and 676 (① in

Fig. 3.14).

The essential amino acid substitutions for changing ligand of NR1 subunit may the substitution to V676 on the lineage to NR1 subunit (② in Fig. 3.14). Valine is a nonpolar amino acid which cannot make a direct hydrogen bond with γ-carboxyl group of glutamate. After the substitution, NR1 might lose the function of binding γ-carboxyl group of glutamate. Therefore, the ligand of NR1 subunit is glycine which loses one carboxyl group comparing to glutamate (Fig. 3.14).

I hypothesized amino acid substitution to A676 on the lineage to NR3A-B is the essential substitution for NR3A-B to lose the function for binding glutamate (③ in Fig. 3.14). Alanine is a nonpolar amino acid which cannot make the direct hydrogen bond to γ-carboxyl group of glutamate.

3.3.2.1.11 Necessary Amino Acid Changes for Ligand Binding Specificity of soluble guanylyl cyclase

Soluble guanylyl cyclase is a receptor for nitric oxide, and is a heterodimer composed of α and β subunits. Soluble guanylyl cyclase composed of $\alpha 1$ and $\beta 1$ or $\alpha 2$ and $\beta 1$ is catalytically active, and active soluble guanylyl cyclase containing $\beta 2$ subunit is not found (Friebe and Koesling 2003). Histidine residue of $\beta 1$ subunit (H105 in bovine) binds the heme iron , and the heme iron is essential for catalysis of nitric oxide (Wedel et al. 1994). The histidine residues of the heme iron binding site are only conserved in $\beta 1$ and $\beta 2$ subunits of soluble guanylyl cyclase, therefore the amino acid substitution to the histidine might occur on the lineage to $\beta 1$ and $\beta 2$ subunits (Fig. 3.10). And I hypothesized this amino acid substitution was the essential for generating nitric oxide sensitive soluble guanylyl cyclase.

3.3.2.2 Phylogenetic Analysis of 1.1 Receptors

3.3.2.2.1 Domain Changes of 1.1 Receptor

1.1 receptor is an ionotropic receptor which is composed of five subunits. Each subunit has four transmembrane regions. According to Pfam database, each subunit of 1.1 receptor has two domains, Neur_chan_LBD and Neur_chan_memb (Fig. 3.5). Neur_chan_LBD domain is the N-terminal ligand binding domain. And the four transmembrane domains are placed on the Neur_chan_memb domain. Fig. 3.15A shows the molecular phylogenetic tree of 1.1 receptors by using Neur_chan_LBD domain. Acetylcholine binding protein has only one domain, Neur_chan_LBD. This protein is not a membrane-bound ionotropic receptor, but a soluble ligand binding protein. I could not define the root point of the phylogenetic tree. I hypothesized two scenarios about the evolution of 1.1 receptors (Fig. 3.15B).

The phylogenetic tree of 1.1 receptors can be hypothesized the left one in Fig. 3.15B, if I think that acetylcholine binding protein is closely related to acetylcholine receptor, and the root point is placed between anion channel receptors and cation channel receptors. This phylogenetic tree (left one in Fig. 3.15B) is also supported by the midpoint rooting. The hypothetical ancestral protein had two domains, Neur_chan_LBD and Neur_chan_memb. The ancestral protein might be ligand binding four transmembrane ionotropic receptor which is similar to 1.1 receptor. Neur_chan_memb domain, the four transmembrane domain may have been lost on the lineage to acetylcholine binding protein. And this domain change might produce soluble acetylcholine binding protein.

If I think the simplest domain structure is the ancestral protein, I hypothesized the phylogenetic tree of 1.1 receptors is the right one in Fig. 3.15B. Acetylcholine

binding protein which is the simplest domain structure can be defined as the out-group. The hypothetical ancestral protein had one ligand binding domain, Neur_chan_LBD. This protein might be similar to soluble acetylcholine ligand binding protein. The insertion of Neur_chan_memb domain which includes four transmembrane regions seems to have produced 1.1 receptor.

I can't decide which phylogenetic tree of 1.1 receptor is plausible or not, but the root point might place around the gray colored region in Fig. 3.15A.

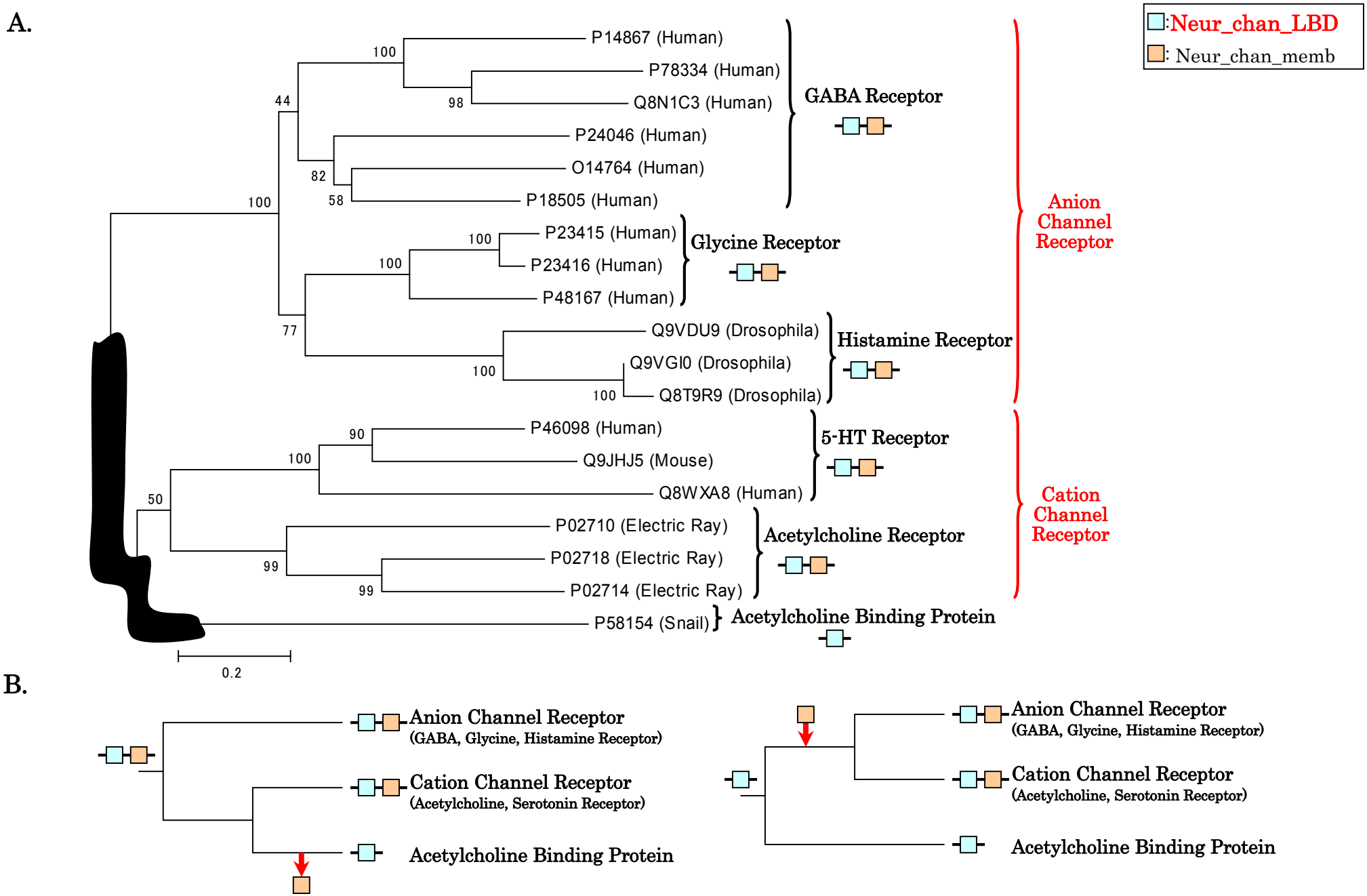


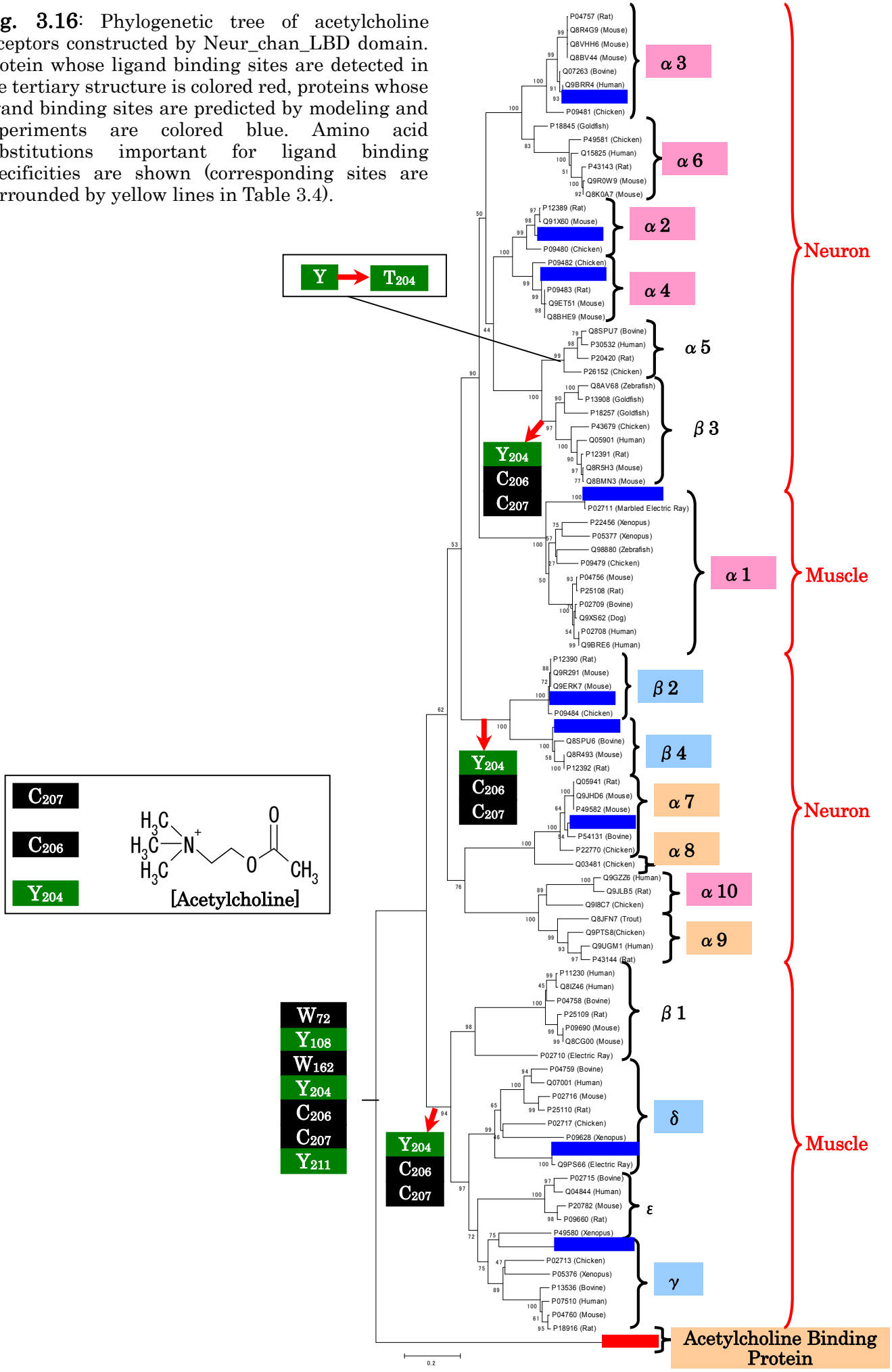
Fig. 3.15: Phylogenetic tree of 1.1 receptor constructed by using Neur_chan_LBD domain

3.3.2.2.2 Necessary Amino Acid Changes for Ligand Binding Specificity of Acetylcholine Receptors

As I described, 1.1 receptor is a membrane-bound ionotropic receptor composed of five subunits. The ligand binding site is placed between subunits. Acetylcholine binding protein which is not a neurotransmitter receptor is a soluble homopentameric protein, and the ligand binding sites placed between subunits, that is, there're five ligand binding sites in one protein (Brejc et al. 2001). Muscle-type acetylcholine receptor is composed of two α1s, β1, γ, and δ subunits in fetal muscles (Mishina et al. 1986) and electrocytes of electric fish(Reynolds and Karlin 1978), and ligand binding sites are placed between α1 and γ, and between α1 and δ (Karlin 2002). Muscle-type acetylcholine receptor is composed of two α1s, β1, ε, δ in adult muscles (Mishina et al. 1986). Neuronal-type acetylcholine receptor is composed of α2-10 and β2-4. α7-9 can form active homopentameric receptor (Cooper et al. 1991; Elgoyhen et al. 1994; Gotti et al. 1994). α2-6 and α10 can form active heteropentameric receptors only when β subunits or other α subunits are co-expressed (Elgoyhen et al. 2001; Le Novere et al. 1996; Ramirez-Latorre et al. 1996; Vernallis et al. 1993), and these subunits combine to produce functional heteropentameric receptors composed of two αs and three βs (Anand et al. 1991) except for α5 and β3 which do not form functional receptors alone or with a single type of α or β subunit (Hogg et al. 2003). Most of receptors are composed of a single type of α subunit and a single type of β subunit, but there are receptors composed of three types of subunit (Hogg et al. 2003). The α5 and β3 subunits form the receptors composed of three types of subunit, when co-expressed with other α or β subunits (Boorman et al. 2000; Groot-Kormelink et al. 1998; Groot-Kormelink et al. 2001;

Ramirez-Latorre et al. 1996). However, the pricise subunit composition of the receptors composed of the three subunit types is not well-known (Hogg et al. 2003).

Fig. 3.16: Phylogenetic tree of acetylcholine receptors constructed by Neur_chan_LBD domain. Protein whose ligand binding sites are detected in the tertiary structure is colored red, proteins whose ligand binding sites are predicted by modeling and experiments are colored blue. Amino acid substitutions important for ligand binding specificities are shown (corresponding sites are surrounded by yellow lines in Table 3.4).



Amino Acid Site Number of Acetylcholine Binding Protein (P58154)		72	73	74	76	100	108	109	– (109A)	123	124	125	131	133	160	162	164	165	183	199	201	204	206	207	211	References
Acetylcholine	α3	W	L	K	I	I	Y	N	N	L	L	K	T	I	G	W+	Y	D	D	K	D/E	Y	C+	C+	Y+	Schapira et al. 2002
	α6	W	L	R	I	I	Y	N	N	L	L	K	T	V	G	W	Y	D	D/E	K	D	Y	C	C	Y	
	α2	W	L	K	E	I	Y	N	N	H	L	F	H/K	V	G	W+	Y	D	D	Y	S	Y	C+	C+	Y+	Schapira et al. 2002
	α4	W	V	K	E	I	Y	N	N	H	L	F	Q/K	T/M	G	W+	Y	D	D	Y	T/S	Y	C+	C+	Y+	Schapira et al. 2002
	α5	W	L	K	E	V/L/I	F/Y	D	N	V	V/I	R/K	T/A	T	G	W	Y	D	D	K	N/S	T	C	C	Y	
	β3	W	L	K/W/Y	E	L/I	F/Y	E	N	I	V	K/R	S/T/Q /M/V	T/M	G	W	Y	D	D	K/R	N/S	D/E/–	F/V/I/–	Y/–/D	Y	
	α1	R	L	K/R	Q	I/V/L	Y+	N	N	L	L	D/E/Q	T/M/I	T/L	G	W+	Y	D	N/T	K	W/S	Y+	C+	C+	Y+	Karlin 2002
	β2	W+	L	T	E	I	Y	N	N	V	V/I	S	F+	L+	R	W	Y	D	D	R	E	D	S	T	Y	Schapira et al. 2002
	β4	W+	L	K	E	V/I	Y	N	N	I/V	V	R	Q/M/L+	L+	R	W	Y	D	D	R	V	Q	–	–	Y	Schapira et al. 2002
	α7	W+	L	Q	S/Y/T	I	Y	N	S	L	V	N	Q+	L+	G	W+	Y	G	S/G	R	E	Y	C+	C+	Y+	Schapira et al. 2002
	α8	W	L	Q	Y	I	Y	N	S	L	V	N	Q	I	G	W	H	S	N	R	E	Y	C	C	Y	
	α10	W	I/V	R	E/A	V	Y	N	K/N	V	L/V	R	R/M	D	G	W	H/Y	G/N	D	R	V	Y	C	C	Y	
	α9	W	I	R	T/S/I	V	Y	N	K	V	L	R	T	D	G	W	Y	N	D	K/R	V	Y	C	C	Y	
	β1	Y/F	L	D/N	E/A	V	L/M	N	N	V/L	V	S/Q	R/S	Q	S/K	Y	Y	D	T/A	L/K	Q/H/W	P/G/V /D	D/Q/ P/–	R/P/–	R/H/Y	
	δ	W+	I/V/M	D/E	A/G/S	V/L	E/Q	N	N	L+	V/I	Y/S/R+	Y/T+	L	S/T	L	Y	T/N/D	G/A	K/V/L	V/I	P/R/ S/D	P/V/L /Y/–	P/L/F /S	R/H/Y	Karlin 2002
	ε	W	I/V	G/Q/E	D/E/ A/Q	V/M	E	N	N	L	V	Y/S	S/Y/T	L	R	Q/K	Y	N/S	A/D	I/K	R/H/L	E/D/H /P/W	G/L/ D/Y	S/T/A	E/Y/F	
	γ	W+	I/V	E	Q	V	E/G	N	N	L+	V	S/Y+	Y+	L	Q	Q	Y	S	A	M/K/R	L/I/Y	V/A/G /H/E	F/L/A /–	P/T	H/Y	Karlin 2002
Acetylcholine Binding Protein		W+	Q	Q+	T	L	Y+	N	–	R+	V+	V+	L+	M+	G	W+	H	H	Y+	K	S	Y+	C+	C+	Y+	Brejc et al. 2001
5-HT	3A	W+	Y+	R+	F/Y+	I	N	E+	F+	Y	V/I	R/H/G	Q	Y	T	W+	H	T	V/I	F/P/V	E/F/T	M/F/L /D/Q	F/I/ S/E/D	S/D	Y/F	Boess et al. 1997; Reeves and Lummis 2002; Steward et al. 2000
	3B	W	Y	R/Q	V	I/L	N	E	F	Y	V	N	R/E	H/Y	N/K	I	H	T	A/S	I/V	Q/S	A/Y	D/I	F/L	–/F	
	3C	W	M	D	V	L	V	E	S	Y	I	S	K	D	S	F	Y	T	V	I	K	P	M	S	Y	
GABA	α2	F	F	R	K	I	H	N	G	R	I	Q	L	T	G	Y	Y	T	Q	T	K	T	E	Y	–	
	α3	F	F	R	T	I	H	N	G	R	L	V	L	T	G	Y	Y	T	Q	I	R	T	E	Y	–	
	α1	F+	F	R+	S+	I	H	N	G	R+	I+	T	L	T	G	Y	Y	T	Q	I	Q	T	E	Y	–	Boileau et al. 2999; Cromer et al. 2002
	α5	F	F	R	S	I	H	N	G	R	L	E	L	T	G	Y	Y	P	Q	N	S	T	E	Y	–	
	α4	F	F	R	T	V	R	N	G	R	I	M	L	T	G	Y	Y	P	Q	T	K	T	E	Y	–	
	α6	F	F	R	T	I	R	N	G	R	I/L	M	L	T	G	Y	Y	P	Q	I/R	S/K	T	E	Y	–	

	ε	I	F	Y/S/H	T	L	R	N	S	L/R	I	H/Y	L	T/S	S	F	Y	D/P	E/Q	I	S/T	V/F	D	F	–	
	γ3	F	F	A	T	I	R	N	S	R	I	W	L	T	S	Y	Y	P	Q	I	T	A	D	Y	–	
	γ4	F	F	A	T	I	R	N	S	R	I	W	L	T	S	Y	Y	P	Q	V	R	A	E	Y	–	
	γ1	I	F	A	T	I	R	N	S	R	I	W	L	T	S	Y	Y	P	Q	I	H	S	D	Y	–	
	γ2	F	F	A	T/M	I	R	N	S	R	I	W	L	T/S	S	Y	Y	P	Q	V/I	K/R	S	D	Y	–	
	ρ1	Y	L	R	Y	I	V	H	S	R	V	Q/Y	L	S	E	Y	Y	T	Q	F	S	T	W	Y	–	
	ρ2	Y	L	R	Y	I	V	H	S	R	V	F	L	S	E	Y	Y	T	Q	F	S	T	W	Y	–	
	ρ3	Y	L	R	Y	I	V	H	S	R	V	H	L	S	E	Y	Y	N	Q	F	S	Y	W	Y	–	
	δ	F	L	H	S	L	V	N	A	R	L	Q	L	S	E	Y	Y	S	Q	N	K	A	Q	F	–	
	π	Y	L	R	R	L	V	E	S	R	L	F	L	A	E	W	Y	D	Q	V/R	Q	T	N	Y	–	
	θ	F	L/F	H	T	L	V/L	N	S	Q	L	H	R/Q	G	E	Y	Y	T	Q	V	F	T	S	Y	–	
	β4	Y	F	Q	S	L	L	N	D	R	L	H	L	G	E	Y	Y	T	Q	E	V	T	S	Y	–	
	β1	Y	F	Q	S	L	L	N	D	R	L	H	L	G	E	Y	Y	T	Q	K	E	T	A	Y	–	
	β2	Y	F	Q	A	L	L	N	D	R	L	H	L	G	E+	Y	Y+	T+	Q	K	V	T+	S	Y+	–	Boileau et al. 2999; Cromer et al. 2002
	β3	Y	F	Q	Y	L	L	N	D	R	L	H	L	G	E	Y	Y	T	Q	N	V	T	A	Y	–	
Glycine	α1	F	L	R	Q	I+	A+	N+	E	R	I	S	L	S	E	F+	Y+	T+	Q	K+	Y+	T+	K	F	–	Schmieden et al. 1993; Vafa 1999; Vandenberg et al. 1992a; Vandenberg et al. 1992b
	α3	F	L	R	K	I	A	N	E	R	I	F	L	S	E	F	Y	T	Q	K	Y	T	K	F	–	
	α2	F	L	R	Q	I	A	N	E	R	I	S	L	S	E	F	Y	T	Q	K	Y	T	K	F	–	
	β	F	L	R	K	L	A	N	E	F	I	F	L	S	E	F	Y	T	Q	K	Y	T	Y	Y	–	
Histamine		F	L/F	A	S/T	I/M	K	N	A	W	L/V	Y	L/F	M	E	L	H	T	Q	T/Q	E/Y	T	N	F	–	

:Basic Amino Acid (Positive Charged)
:Acidic Amino Acid (Negative Charged)
:Uncharged Polar Amino Acid
:Nonpolar Amino Acid

** Ligand binding site which interacts directly with ligand detected by tertiary structure study.
+ Ligand binding site which interacts with ligand predicted by modeling and experiments.

Table 3.4: Important sites for ligand binding in 2.1 receptors

The ligand binding sites of 1.1 receptors are placed in the interface of α subunit and its adjacent subunit (Corringer et al. 2000). It is supposed that two acetyocholine binding sites in muscle type receptor are between α and β, and α and δ (or ε), two binding sites in neuronal heteropentameric receptor composed of two α s and three β s are between α and β, and five binding sites in neuronal homopentameric receptor are between α and another α (Hogg et al. 2003; Karlin 2002). Previous studies demonstrated that agonists mainly bind to α subunit (Middleton and Cohen (1991); Oswald RE and Changeux (1982); Reynolds and Karlin (1978)). This suggests that α subunit contributes the principal component of the binding site, and the adjacent subunit makes up the complementary component (Corringer et al. 1995). The subunits which have both the principal and complementary components of the lingand binding sites are colored orange in Fig. 3.16. The subunits which have the principal component and the subunits which have the complementary component are colored pink and light blue in Fig. 3.16, respectively. Fig. 3.16 is the phylogenetic tree constructed by using Neur_chan_LBD domain, the ligand binding domain. I defined acethlcholine binding protein as the out-group. From this tree, the ancestral protein of the acetylcholine receptor seems to have owned both the principal and complementary components. But through the evolution, some subunits lost the principal or complementary component.

Amino acid sites which are important for ligand binding are shown in Table 3.4. These sites are detected by tertiary structure studies, computer modeling studies and mutation studies (bibliographies cited are listed in Table 3.4). Colors of amino acid sites in Fig. 3.17 are written in Table 3.4. Amino acid numbers I'll describe in this chapter about the 1.1 receptors correspond to the acetylcholine binding protein of snail (P58154)

(Table 3.4). If I think parsimoniously, the ligand binding sites which are conserved from the ancestral protein of the acetylcholine receptor seem to be W72, Y108, W162, Y204, C206, C207 and Y211 (Fig. 3.17). Among these ligand binding sites, Y204, C206 and C207 are conserved in all subunits which have the principal component, but lost in all subunits which have the complementary component. These three amino acids (Y204, C206 and C207) may interact with acetyl group of acetylcholine by van der Waals force (Schapira et al. 2002). And C206 and C207 makes disulfide bond and form ligand binding pore (Brejc et al. 2001; Karlin 2002; Schapira et al. 2002). I hypothesized the points where the amino acid substitutions from the three amino acids to others had occurred. Red arrows show the substitutions. All the substitutions might occur on the lineages to subunits which lost the principal component. Threfore, I hypothesized that the amino acid substitutions from Y204, C206 and C207 to others had made subunits lose the formation of the principal component. From tertiary structure and computer modeling analysis (Brejc et al. 2001; Schapira et al. 2002), these three amino acids are placed near the acetyl group of acetylcholine and may interact with the acetyl group. The amino acid substitutions from Y204, C206 and C207 to others may lose the function for binding the acetyl group.

3.3.2.2.3 Necessary Amino Acid Changes for Ligand Binding Specificity of 5-HT_3 Receptors

5-HT_3 receptors are pentameric receptors, and there are three kinds of subunits; 3A, 3B and 3C (Reeves and Lummis 2002). 3A can form functional homopentameric receptor, and this subunit has the ligand binding sites (Reeves and Lummis 2002). 3B can not form functional homopentameric receptor, but

heteropentameric receptor composed of 3A and 3B is functional receptor (Hanna et al. 2000). The main effect of 3B seems to be biophysical properties rather than pharmacological properties (Brady et al. 2001). Therefore, 3B may not bind ligand. Functions of 3C are unknown, 3C may works for regulation of the receptors (Reeves and Lummis 2002).

The important sites for binding 5-HT are shown in Table 3.4. If I think parsimoniously referring to the phylogenetic tree of 1.1 receptors (Fig. 3.15) and the tree of 5-HT_3 receptors (Fig. 3.17), I hypothesized that the six amino acid substitutions at the positions of 72, 109, 162, 204, 206 and 207 are essential for ligand binding specificity of 5-HT_3 receptors (Fig. 3.17). In the last section, I predicted Y204, C206 and C207 are important for the formation of the (+) ligand binding site. These amino acids are not conserved in 5-HT_3 receptors. Therefore, I predicted the amino acid substitutions from Y204, C206 and C207 to others might occur on the lineage to 5-HT_3 receptors (Fig. 3.17), and these amino acid substitutions may make 1.1 receptors lose the function of binding acetylcholine. E109 is predicted to be placed near the amino group of 5-HT from the modeling study (Boess et al. 1997). I hypothesized that E109 could interact with the amino group. The amino acid substitution to E109 might occur on the lineage to 5-HT_3 receptor, and the substitution might produce the interaction between 5-HT_3 receptor and the amino group of 5-HT (Fig. 3.17). I hypothesized W72 and W162, two aromatic amino acids, which formed aromatic pore in the ligand binding sites had conserved from the ancestral protein of the 5-HT_3 receptors. Although W72 is conserved among the all 5-HT_3 receptors, W162 is only conserved among the 5-HT_{3A} receptors. I hypothesized that the amino acid substitution from W162 to I on the lineage to 5-HT_{3B} and the amino acid substitution from W162 to F on the lineage to 5-HT_{3C} made 5-HT_3 receptors lose

the 5-HT binding function.

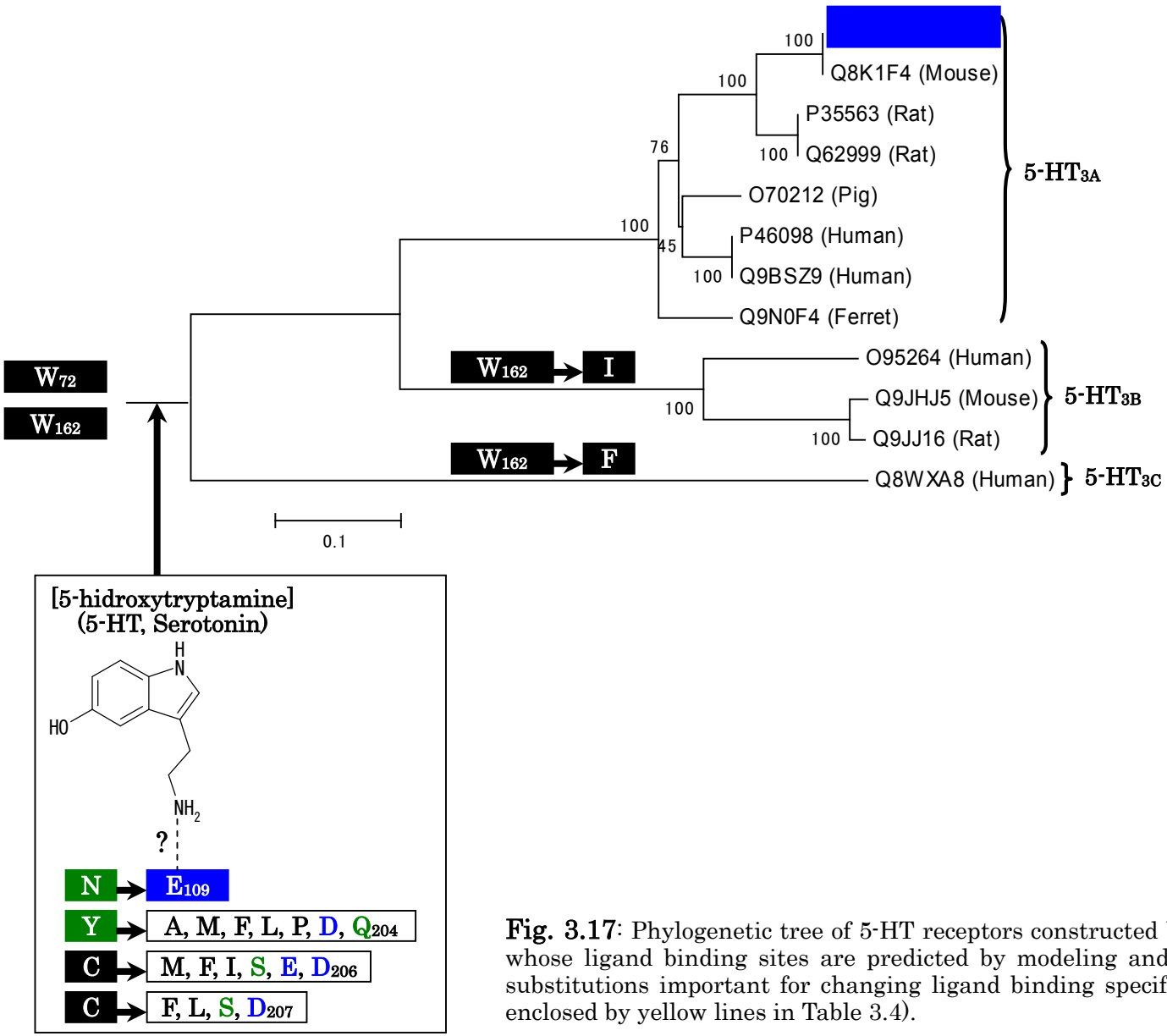


Fig. 3.17: Phylogenetic tree of 5-HT receptors constructed by using Neur_chan_LBD domain. Protein whose ligand binding sites are predicted by modeling and experiments is colored blue. Amino acid substitutions important for changing ligand binding specificities are shown (corresponding sites are enclosed by yellow lines in Table 3.4).

3.3.2.2.4 Necessary Amino Acid Changes for Ligand Binding Specificity of $GABA_A$ Receptors

$GABA_A$ receptors are pentameric receptor, and subunits of $GABA_A$ receptor are categorized into α 1-6, β 1-4, γ 1-4, δ, ε, π and ρ 1-3 (Watanabe et al. 2002). Some researchers define GABA receptors composed of ρ 1-3 subunits as $GABA_C$ receptor, because the pharmacological features of the $GABA_C$ receptor are different from those of $GABA_A$ receptor (Bormann 2000). But according to IUPHAR (International Union of Pharmacology) definition, ρ 1-3 subunits are $GABA_A$ receptor subunits (Barnard et al. 1998). At least each one α, β and γ subunits are included in the functional receptor, and most subunit compositions of $GABA_A$ receptors are $2\alpha 2\beta 1\gamma$, $2\alpha 1\beta 2\gamma$, $1\alpha 2\beta 2\gamma$ (Watanabe et al. 2002). γ can be replaced by δ. The ligand binding sites of $GABA_A$ receptor are placed between α and β subunits (Watanabe et al. 2002). The ligand binding sites of $GABA_A$ receptor are predicted to be F72, R74, R123 and I or L124 in α subunit and E160, Y164, T165, T204 and Y207 in β subunit from computer modeling and mutation studies (Boileau et al. 2999; Cromer et al. 2002) (Table 3.4). Most of these amino acids are conserved in both α subunits and β subunits. Fig. 3.18 is the phylogenetic tree of $GABA_A$ receptors. If I think parsimoniously, the ancestral protein of $GABA_A$ receptors might have predicted ligand binding sites of both α subunit and β subunit (F72, R74, R123, I or V or L124, E160, Y164, T165, T204 and Y207) (Fig. 3.18). Comparing ligand binding sites between α and β subunits, the amino acid sites at the positions of 72, 74 and 160 are different between the two subunits. I hypothesized these two differences are the cause for the different ligand binding feature between α and β subunits. If I think parsimoniously, the amino acid substitutions at the position of 160 (from E to G) and at the positions of

72 (from F to Y) and 74 (from R to Q) might occur on the lineage to α and β subunit respectively. Although there are no evidences, I hypothesized E160 could interact with carboxyl group of GABA. Amino acid substitution from E160 to G might make $GABA_A$ receptor weaken the interaction with the carboxyl group of GABA.

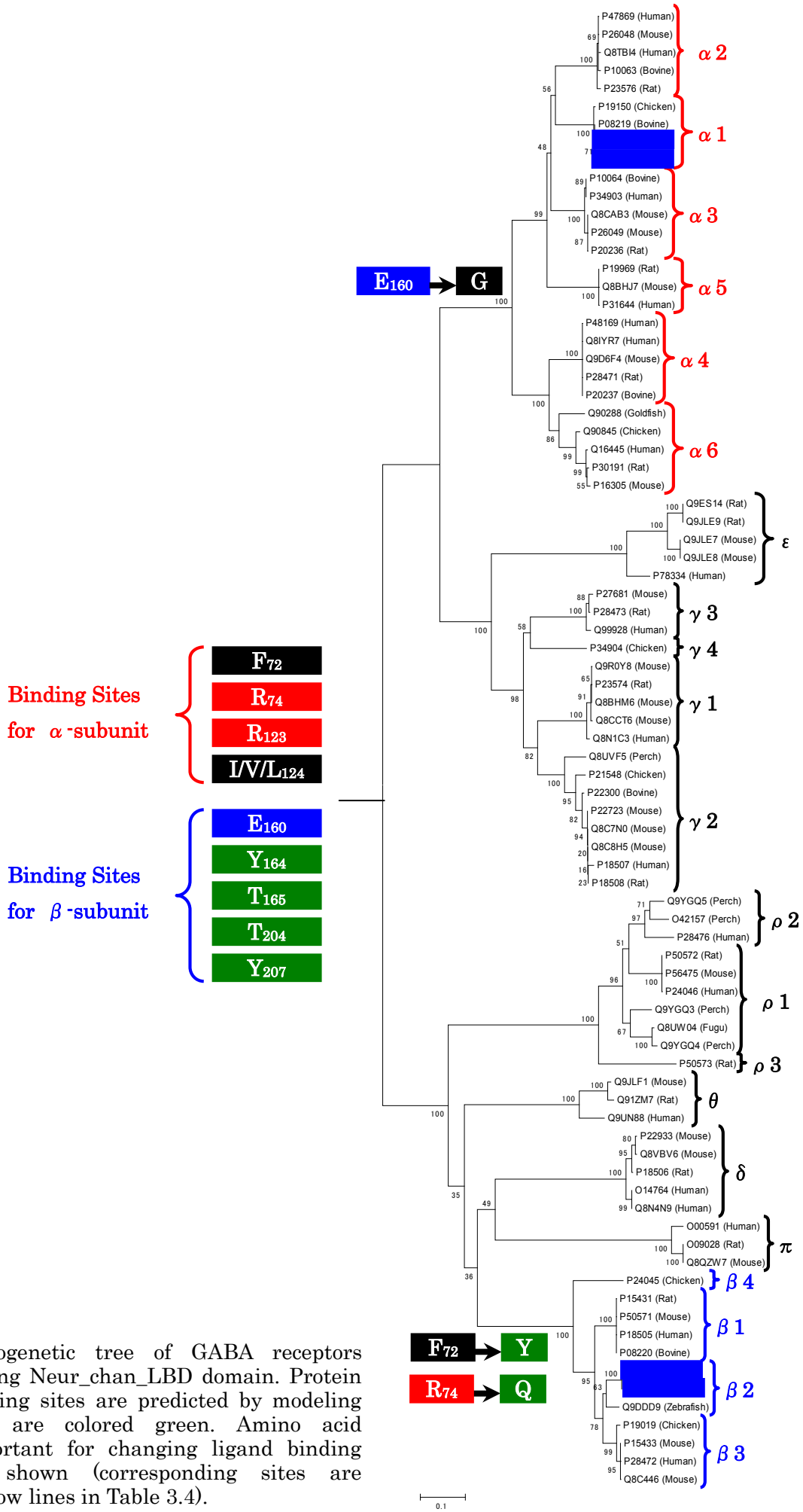


Fig. 3.18: Phylogenetic tree of GABA receptors constructed by using Neur_chan_LBD domain. Protein whose ligand binding sites are predicted by modeling and experiments are colored green. Amino acid substitutions important for changing ligand binding specificities are shown (corresponding sites are surrounded by yellow lines in Table 3.4).

3.3.2.2.5 Necessary Amino Acid Changes for Ligand Binding Specificity of Glycine Receptors

Glycine receptor is composed of α and β subunits, which form pentameric structure, and cytoplasmic sutunit. Although α and β subunits are homologous to other 1.1 receptor subunits, cytoplasmic subunit is not homologous. Fig. 3.19 shows the phylogenetic tree of α and β subunits. α subunits can form functional homopentameric structure, but β subunits can not (Jentsch et al. 2002).

The important sites for ligand binding in α subunit are shown in Table 3.4. I predicted amino acid changes of these sites by thinking parsimoniously referring to Table 3.4, Fig. 3.15 and Fig. 3.19. Among these sites, I hypothesized that I121, N130, Y189, T190 and T232 had been conserved from the ancestral protein, and amino acid substitutions to A129, F187, K228 and Y230 occurred on the lineage to α and β subunits (Fig. 3.19). Only one amino acid site is different between α and β subunits, and the site is the position of 100 (I in α subunit and L in β subunit). I hypothesized that the amino acid substitution from I100 to L might have occurred on the lineage to β subunit (Fig. 3.19). But I don't know whether this substitution is the essential substitution for loosing the function for glycine binding in β subunit.

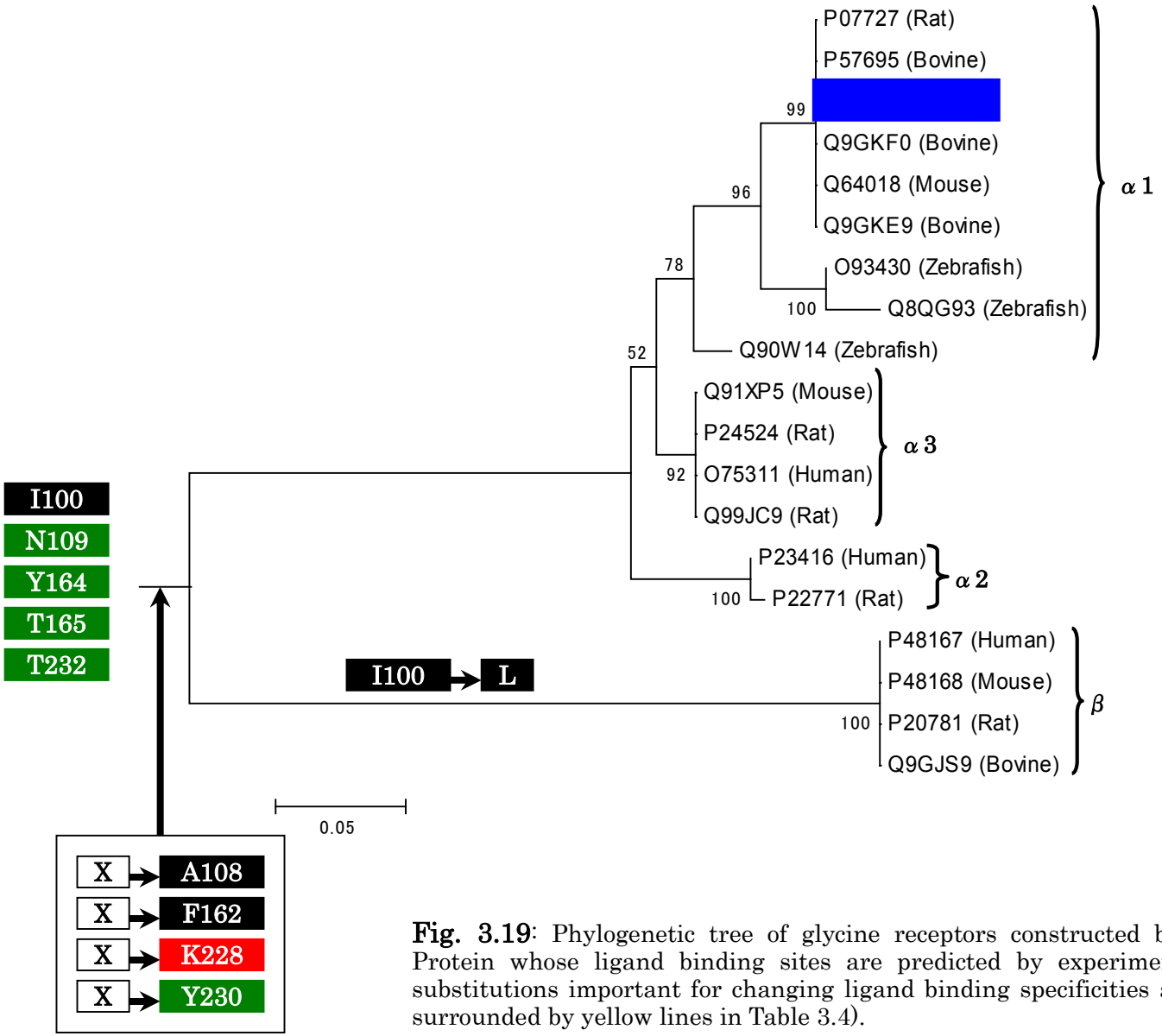


Fig. 3.19: Phylogenetic tree of glycine receptors constructed by using Neur_chan_LBD domain. Protein whose ligand binding sites are predicted by experiments are colored blue. Amino acid substitutions important for changing ligand binding specificities are shown (corresponding sites are surrounded by yellow lines in Table 3.4).

3.3.2.3 Phylogenetic Analysis of 1.4 Receptors

1.4 receptors are all P2X purinoceptors. They have two transmembrane regions, and their ligand is ATP. P2X receptors are homo- or heteropolymers, and subunits are categorized into $P2X_1$-$P2X_7$ (Khakh et al. 2001; North 2002). Fig. 3.20 shows the phylogenetic tree of P2X receptors constructed by using P2X_receptor domain. All subunits have P2X_receptor domain which includes the ligand binding sites and the two transmembrane regions. I could not define the root of the phylogenetic tree of 1.4 receptors. There are no reports of sequences homologous with P2X receptors in invertebrates, therefore these receptors might be generated after the divergence between vertebrates and invertebrates.

Putative ligand binding sites of P2X receptors are shown in Fig. 3.21. These four sites were predicted by mutational studies of human $P2X_1$ receptor (Ennion et al. 2000). These four amino acids are conserved all of P2X receptors. Therefore, the ancestral protein might have the four amino acids and the ATP binding function.

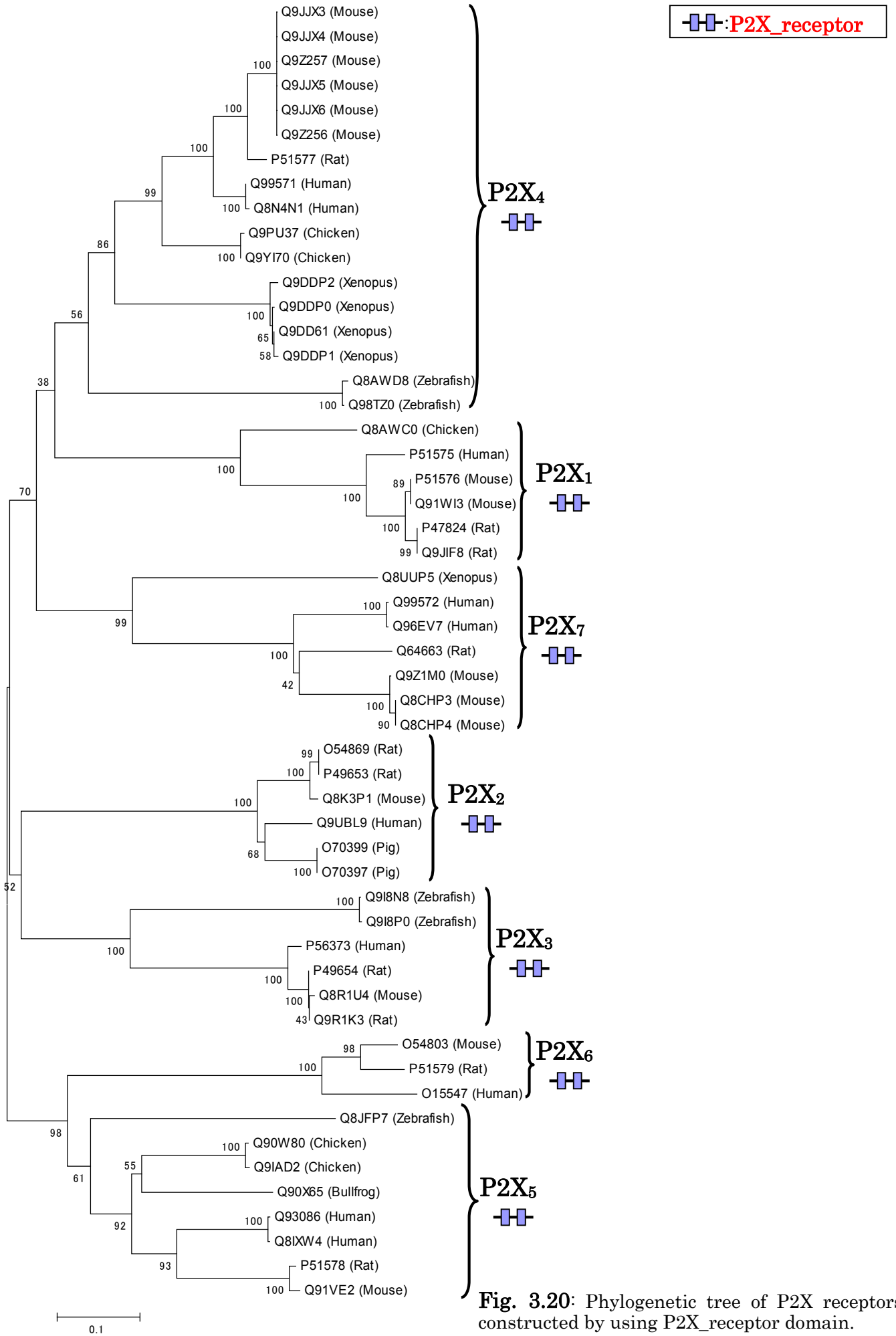


Fig. 3.20: Phylogenetic tree of P2X receptors constructed by using P2X_receptor domain.

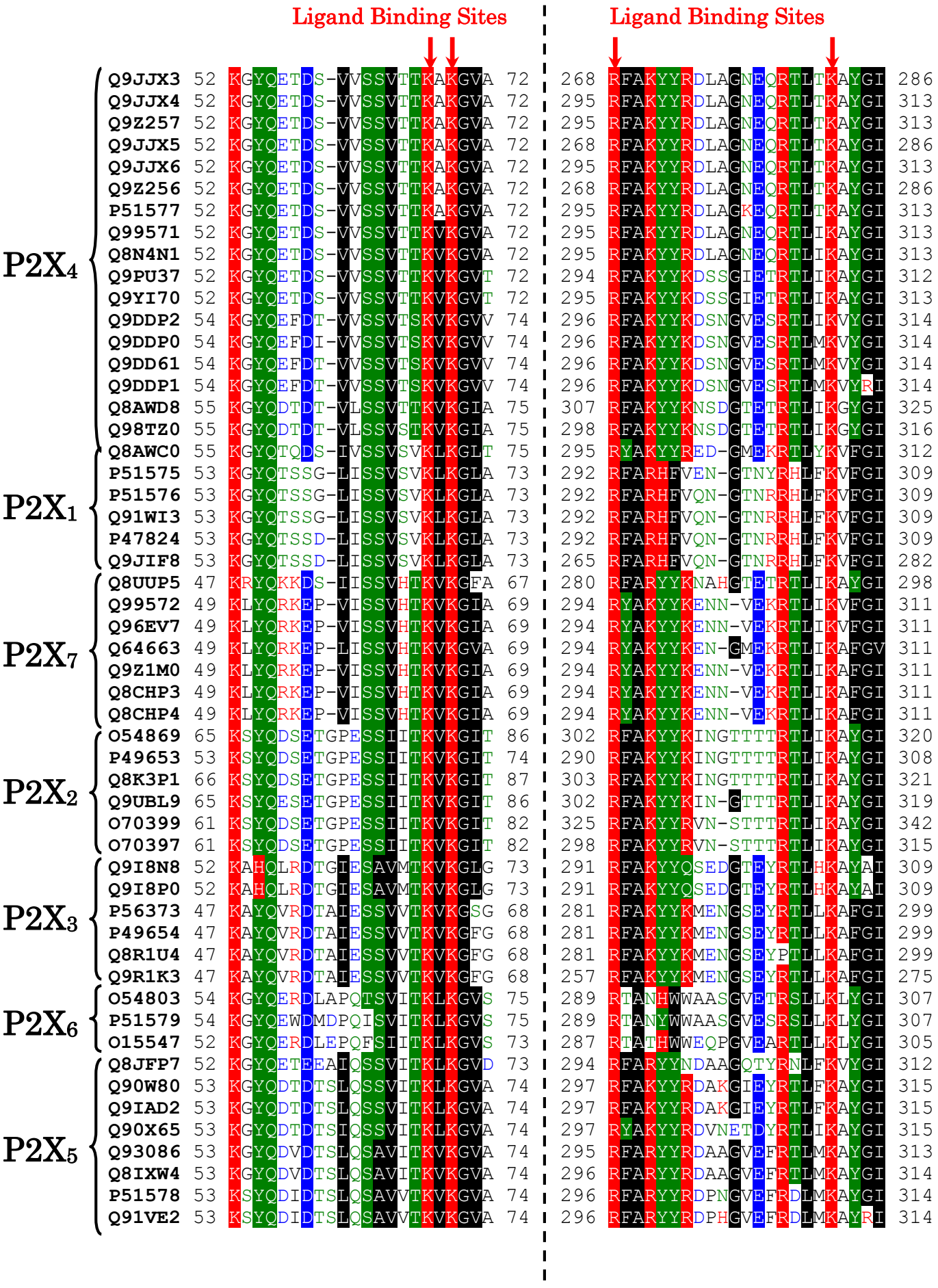


Fig. 3.21: Multiple alignment of P2X receptors. Red arrows indicate putative ligand binding sites.

3.3.2.4 Phylogenetic Analysis of 2.1 Receptors

2.1 receptors are GPCR class A, and all of 2.1 receptors exist in eumetazoa. They have seven transmembrane regions. There are many other receptors for neurotransmissions not listed in Table 3.1. They are opsin, olfactory receptor, neuropeptide receptor and so on (Fig. 3.22). Fig. 3.22 is the phylogenetic tree of 2.1 receptors. Many of bootstrap values in the tree are very low, therefore this tree is not necessarily reliable. Because it was difficult to define the root of the tree, I inferred the root by mid-point rooting. The ancestral protein might have 7tm_1 domain including seven transmembrane regions. Domain changes might occur on the lineage to glycoprotein hormone receptor and relaxin receptor. Amine receptors make a cluster, and P2Y purinoceptors get together. Ancestral receptors of amine receptors and purinoceptors seem to have been peptide receptors. I did further phylogenetic analysis of amine receptors which are listed in Table 3.1 and clustered in the molecular phylogenetic tree (Fig. 3.22).

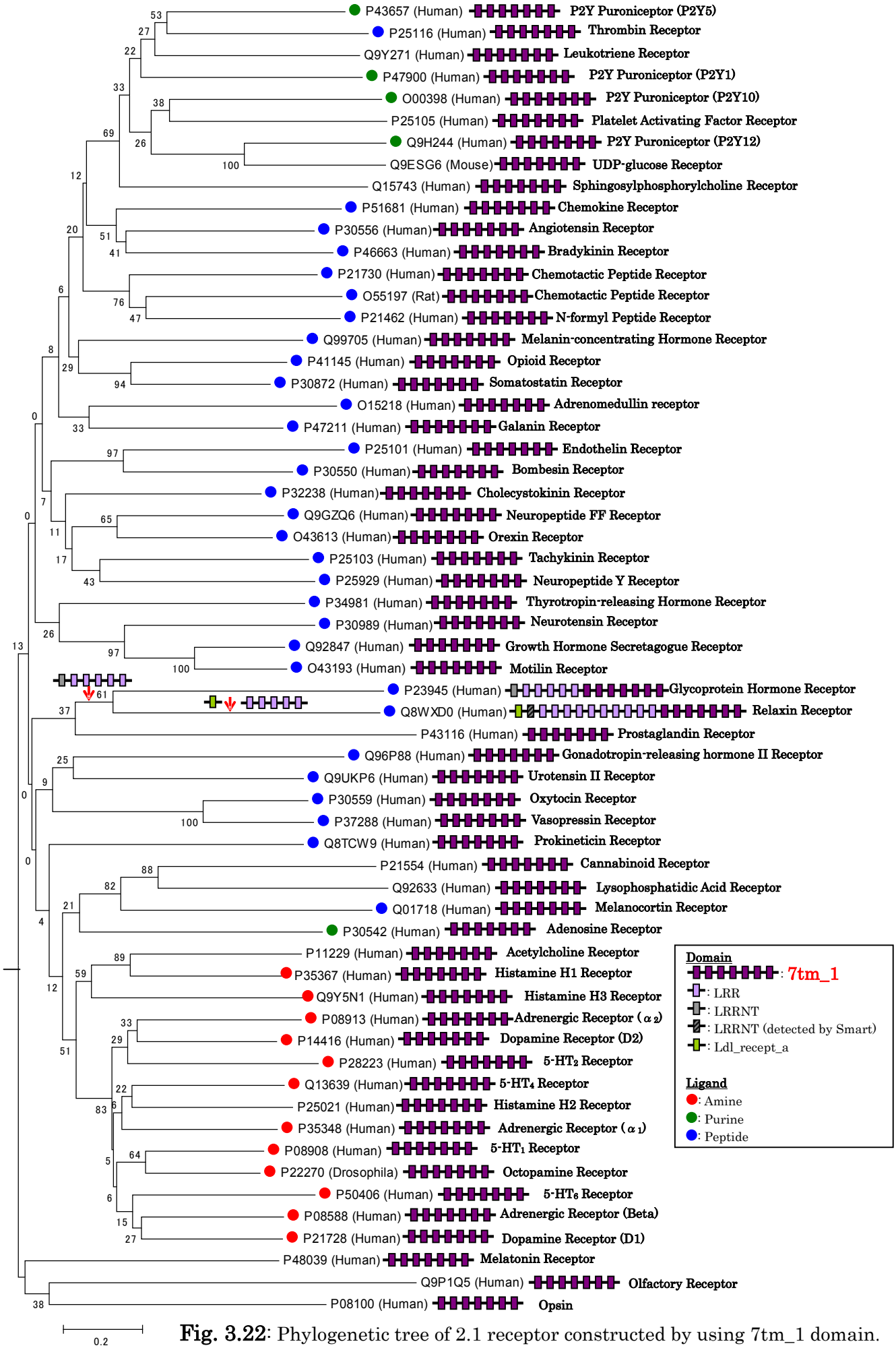


Fig. 3.22: Phylogenetic tree of 2.1 receptor constructed by using 7tm_1 domain.

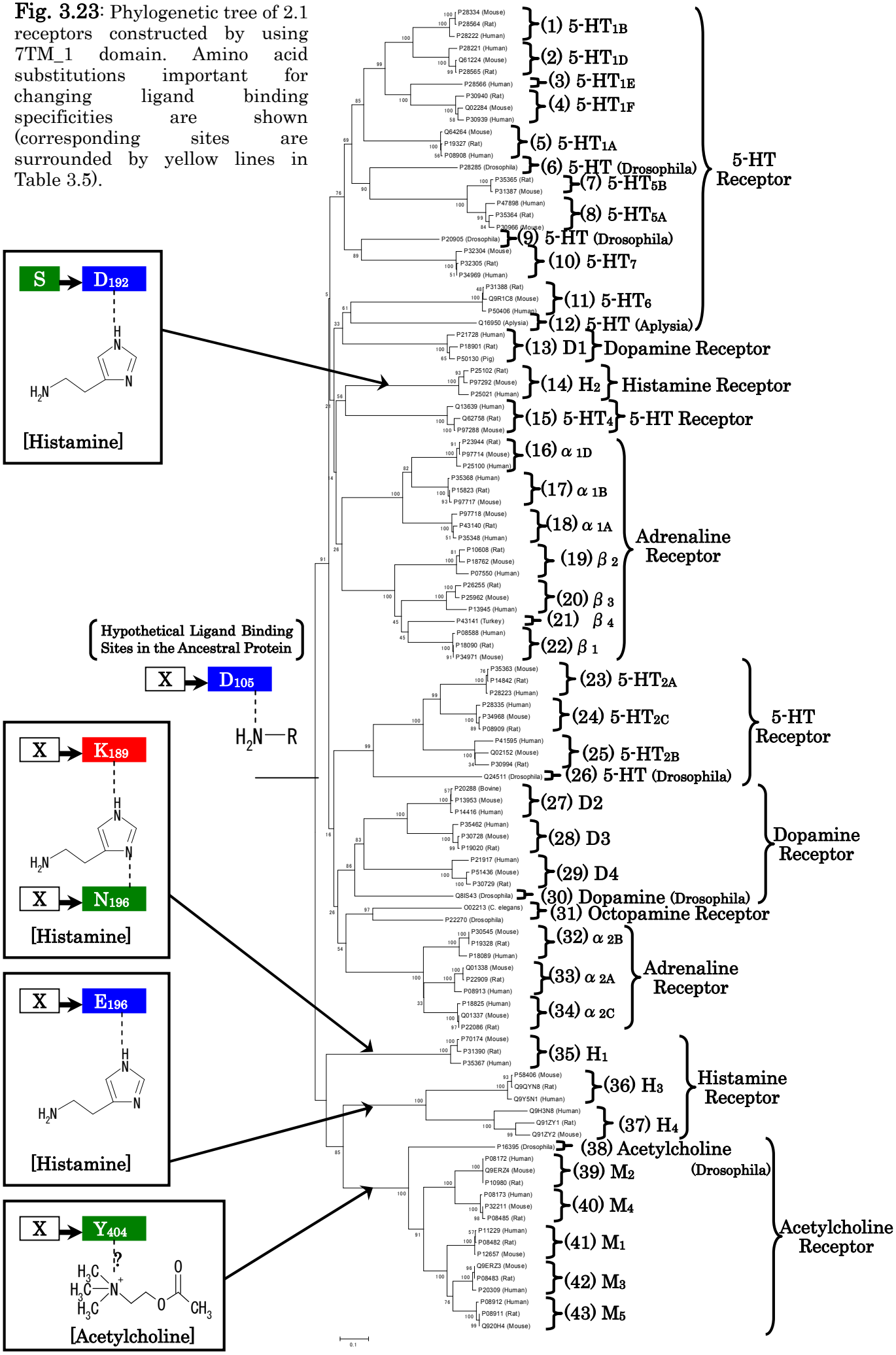
Necessary Amino Acid Changes for Ligand Binding Specificity of Amine Receptors

Fig. 3.23 is the phylogenetic tree of amine and acetylcholine receptors which make a cluster in Fig. 3.22. I predicted the essential amino acid substitutions for generating each receptor. The ligand binding sites of the receptors are shown in Table 3.5. Amino acid site numbers I describe in this section are site numbers of acetylcholine M_1 receptors in rat (P08482). D105 is conserved among all of amine and acetylcholine receptors. D105 is predicted to interact with amino group in amine (Boess et al. 1998; Cavalli et al. 1996; Elz et al. 2000; Fukui et al. 1999; Gantz et al. 1992; Manivet et al. 2002; Mialet et al. 2000; Ohta et al. 1994; Seeber et al. 2003; Strader et al. 1988; Roth et al. 1998; Uveges et al. 2002; Vaidehi et al. 2002; Wang et al. 1991) or acetyl group in acetylcholine (Page et al. 1995). D105 is conserved not only in amine and acetylcholine receptors, but also in opioid, somatostatin and urotensin II receptors. But the lineages of opioid, somatostatin and urotensin II receptors are apart from the lineage of amine and acetylcholine receptors. Therefore, amino acid substitution to D105 seems to have occurred on the lineage to amine and acetylcholine receptors, if I think parsimoniously. And this substitution might make molecules which have amine or acetyl group possible to bind to 2.1 receptors, therefore I hypothesized the amino acid substitution to D105 was the essential amino acid substitution for generating amine and acetylcholine receptors.

I predicted the essential amino acid substitutions for generating histamine receptor and acethlcholine receptors (Fig. 3.23). There are four subtypes in histamine receptor, H_1-H_4. The essential substitutions for generating H_2 receptors, H_1 receptors and H_3 and H_4 receptors are the substitution from S to D192, the substitutions to K189 and N196 and the substitution to E196, if I think parsimoniously. And these

substitutions might make imidazole ring of histamine possible to bind 2.1 receptors. Although H3 and H4 seem to have generated by the same substitutions, different substitutions might produce H1, H2 and H3-4 receptors independently. The amino acid substitution to Y404 is the only substitution on the lineage to acetylcholine receptor which produces amino acids which may interact with acetylcholine. From modeling study (Lu et al. 2001), Y404 may be possible to interact with acetyl group of acetylcholine. But D105, which may have conserved from the ancestor of amine and acetylcholine receptors, is also possible to interact with acetyl group of acetylcholine. Therefore, I can't decide whether the amino acid substitution to Y404 is the essential amino acid substitution or not. I also tried to predict other essential amino acid substitutions for generating each receptor, but it is impossible to predict those substitutions. Because many amino acid substitutions might occur in relatively very short time, and I couldn't trace the evolutionary history of amine receptors. Moreover, bootstrap values of the phylogenetic tree in Fig. 23 are low in amine receptor lineages. Therefore, I could not predict exact traces of amine receptor evolution.

Fig. 3.23: Phylogenetic tree of 2.1 receptors constructed by using 7TM_1 domain. Amino acid substitutions important for changing ligand binding specificities are shown (corresponding sites are surrounded by yellow lines in Table 3.5).



Amino Acid Site Number of Acetylcholine M_1 Receptor (P08482)		101	105	106	108	109	150	188	189	192	193	196	378	381	382	385	403	404	405	408	References
1	5-HT_{1B} Receptor	W	D	I	C	C	W	Y	T	S	T	A	W	F	F	S	F	N/T	W	Y	
2	5-HT_{1D} Receptor	W	D	I	C	C	W	Y	T	S	T	A	W	F	F	S/T	F	T	W	Y	
3	5-HT_{1E} Receptor	W	D	M	C	C	W	Y	T	S	T	A	W	F	F	E	L	T	W	Y	
4	5-HT_{1F} Receptor	W	D	I	C	C	W	S	T	S	T	A	W	F	F	E	L	A/T	W	Y	
5	5-HT_{1A} Receptor	F	D$^{\#\#}$	V/M	C	C	W	Y$^{\#\#}$	T	S$^{\#\#}$	T$^{\#\#}$	A	W	F$^{\#\#}$	F	A	I	N/T	W	Y	Seeber et al. 2003
6	5HT-Receptor (Drosophila)	W	D	V	C	C	W	Y	Q	A	T	T	W	F	F	A	F	L	W	Y	
7	5-HT_{5B} Receptor	W	D	V	C	C	W	Y	A	S	T	A	W	F	F	E	F	L	W	Y	
8	5-HT_{5A} Receptor	W	D	V	C	C	W	Y	T/A	S	T	A	W	F	F	E	F	L	W	Y	
9	5HT-Receptor (Drosophila)	W	D	V	C	C	W	Y	Q	A	T	S	W	F	F	A	F	L	W	Y	
10	5-HT_7 Receptor	F	D	V	C	C	W	Y	T	S	T	A	W	F	F	S	C/F/L	L	W	Y	
11	5-HT_6 Receptor	W	D^{+}	V	C	C	W	F/Y	V	A	S	T	W	F	F	N/S	L	T	W	Y	Boess et al. 1998
12	5-HT Receptor (Aplysia)	W	D	V	C	E	W	Y	S	Y	S	C	W	F	F	N	T	I	W	Y	
13	Dopamine D1 Receptor	W	D	I	C	S	W	Y	A	S	S	S	W	F	F	N	F	V	W	W	
14	Histamine H_2 Receptor	Y	D^{+}	V	L	C	W	Y$^{\#\#}$	G	D$^{\#\#,+}$	G	T^{+}	W	Y	F	F	V	L	W	Y	Gantz et al. 1992; Nederkoorn et al. 1996
15	5-HT_4 Receptor	R	D$^{\#\#}$	V	L	T	W	Y	A	C	S$^{\#\#}$	A	W	F	F	N	F	L	W	Y	Mialet et al. 2000
16	Adrenaline α_{1D} Receptor	W	D	V	C	C	W	Y	A	S	S	S	W	F	F	L	I	F	W	Y	
17	Adrenaline α_{1B} Receptor	W	D$^{\#\#}$	V	C	C	W	Y	A	S$^{\#\#}$	S$^{\#\#}$	S$^{\#\#}$	W	F	F	L	V	F	W	Y	Cavalli et al. 1996
18	Adrenaline α_{1A} Receptor	W	D	V	C	C	W	Y	V/A	S$^{\#\#,+}$	A	S^{+}	W	F	F	M/L	V/A/T	F	W	Y	Hwa and Perez 1996; Perez et al. 1998
19	Adrenaline β_2 Receptor	W	D^{+}	V	C	V	W	Y	A	S	S^{+}	S^{+}	W	F	F	N	L	N	W	Y	Strader et al. 1988; Strader et al. 1989
20	Adrenaline β_3 Receptor	W	D	V	C	V	W	Y	A/V	S	S	S	W	F	F	N	L	N	W	Y	
21	Adrenaline β_4 Receptor	W	D	V	C	V	W	Y	A	S	S	S	W	F	F	N	L	N	W	Y	
22	Adrenaline β_1 Receptor	W	D$^{\#\#}$	V	C	V	W	Y	A	S	S$^{\#\#}$	S$^{\#\#}$	W	F	F$^{\#\#}$	N	F/L	N	W	Y	Vaidehi et al. 2002
23	5-HT_{2A} Receptor	W	D^{+}	V	F	S^{+}	W^{+}	F	V	G	S	A/S	W^{+}	F	F^{+}	N	F	V	W^{+}	Y	Roth et al. 1998
24	5-HT_{2C} Receptor	W	D^{+}	V	F	S^{+}	W	F	V	G	S	A	W	F	F^{+}	N	F	V	W	Y	Roth et al. 1998
25	5-HT_{2B} Receptor	W$^{\#\#}$	D$^{\#\#}$	V	F$^{\#\#}$	S$^{\#\#}$	W	F	M/I	G	S	A	W	F	F$^{\#\#}$	N$^{\#\#}$	F$^{\#\#}$	V/S	W	Y	Manivet et al. 2002
26	5-HT Receptor (Drosophila)	Y	D	V	A	C	W	F	F	G	S	A	W	F	F	N	C	L	W	Y	
27	Dopamine D2 Receptor	F	D	V	M	C	W	F	V	S^{+}	S^{+}	S	W	F	F	H	F	T	W	Y	Coley et al. 2000; Cox et al. 1992
28	Dopamine D3 Receptor	F	D	V	M	C	W	F	V	S^{+}	S	S	W	F	F	H	T/F	T	W	Y	Sartania and Strange 1999
29	Dopamine D4 Receptor	L	D	V	L	C	W	Y	V	S	S	S	W	F	F	H	V	T	W	Y	
30	Dopamine Receptor (Drosophila)	Y	D	V	C	S	W	F	I	S	S	S	W	F	F	N	T	T	W	Y	
31	Octopamine Receptor	W/F	D	I/V	C/L	C	W	F/Y	V	S	S/A	S	W	F	F	Y	I/F	T	W	Y	
32	Adrenaline α_{2B} Receptor	Y	D	V	F	C	W	Y	I	S	S/T	S	W	F	F	Y	F	F	W	Y	
33	Adrenaline α_{2A} Receptor	Y	D^{+}	V	F	C	W	Y	V/I	S	C/S	S^{+}	W	F	F	Y	F	F	W	Y	Wang et al. 1991
34	Adrenaline α_{2C} Receptor	Y	D	V	F	C	W	Y	I	S	C	S	W	F	F	Y	F	F	W	Y	
35	Histamine H_1 Receptor	W	D$^{\#\#,+}$	Y	A	S	W	F	K$^{\#\#,+}$	T	A	N$^{\#\#,+}$	W^{+}	Y^{+}	F	F	T	I	W	Y	Elz et al. 2000; Fukui et al. 1999; Leurs et al. 1995; Ohta et al. 1994; ter Laak et al. 1995; Wieland et al. 1999
36	Histamine H_3 Receptor	W	D$^{\#\#}$	Y	L	C	W	F	L	A	S	E$^{\#\#}$	W	Y	T	M	S	F	W	W	Uveges et al. 2002
37	Histamine H_4 Receptor	W	D	Y	L	C	W	I/V/F	L/A	T	S/L/A/M	E	W	Y	S/C	T	A	F	W	W	
38	Acetylcholine Receptor (Drosophila)	W	D	Y	A	S	W	I	T	T	A	A	W	Y	N	V	F	Y	A	Y	
39	Acetylcholine M_2 Receptor	W	D	Y	V	S	W	V	T	T	A	A	W	Y	N	V	G	Y	W	Y	
40	Acetylcholine M_4 Receptor	W	D	Y	V	S	W	V	T	T	A	A	W	Y	N	V	G	Y	W	Y	
41	Acetylcholine M_1 Receptor	W	D$^{\#\#,+}$	Y$^{\#\#}$	A	S	W	I	T	T	A	A	W	Y$^{\#\#,+}$	N	V	G	Y$^{\#\#}$	W	Y$^{\#\#}$	Page et al. 1995; Lu et al. 2001; Ward et al. 1999
42	Acetylcholine M_3 Receptor	W	D	Y	A	S	W	I	T	T	A	A	W	Y	N	V	G	Y	W	Y	
43	Acetylcholine M_5 Receptor	W	D	Y	A	S	W	I	T	T	A	A	W	Y	N	V	G	Y	W	Y	

: Basic Amino Acid (Positive Charged)
: Acidic Amino Acid (Negative Charged)
: Uncharged Polar Amino Acid
: Nonpolar Amino Acid

Ligand binding site which interacts directly with ligand predicted by computer modeling.
\+ Ligand binding site which interacts with ligand predicted by mutation study.

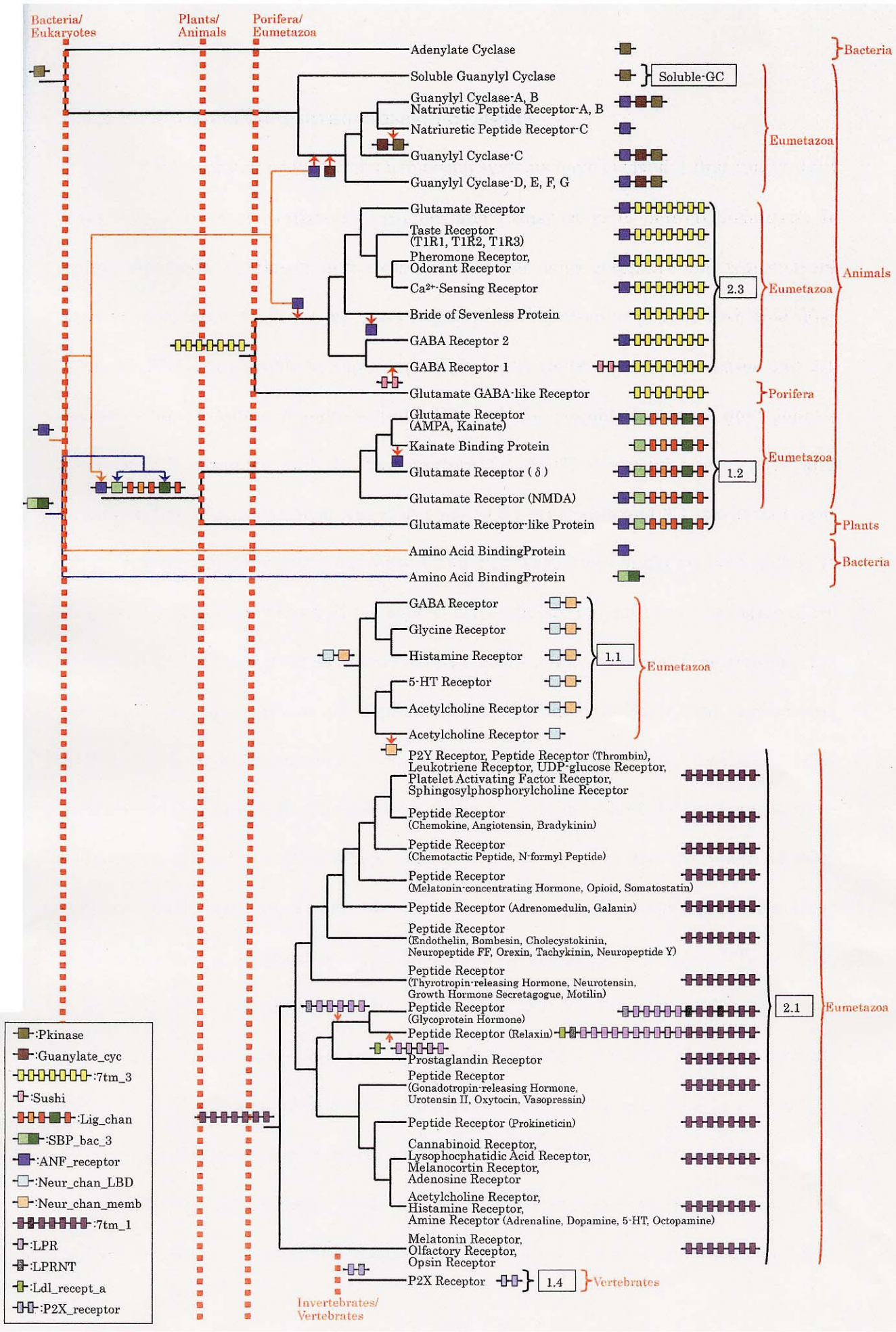
Table 3.5: Important sites for ligand binding in 2.1 receptors

3.4 Conclusion

3.4.1 Evolution of Neurotransmitter Receptors

I did phylogenetic analysis of neurotransmitter receptors and investigated how amino acid substitutions had changed protein functions through evolution. Fig. 3.24 is the summarized phylogenetic trees of receptors I analyzed. As I explained in previous sections, I hypothesized that some domains in 1.2 and 2.3 receptors and soluble guanylyl cyclase had been generated before the divergence between bacteria and eukaryotes. And these receptors may have been generated and diversified after the divergence after bacteria and eukaryotes. From my phylogenetic analysis, 1.1 receptors and 2.1 receptors may have been produced after the divergence between porifera and eumetazoa. Because 1.4 receptors exist only in vertebrates, I hypothesized that 1.4 receptor was generated after the divergence between invertebrates and vertebrates. If I trace the history of receptors by domain unit, 1.1, 1.4 and 2.1 receptors are relatively new. Although 2.1 receptor is relatively new one, this receptor may have diversified rapidly after the divergence between porifera and eumetazoa. New functions may have been acquired by domain changes in the history of neurotransmitter receptors, but ligand specificities may have changed a small number of amino acid substitutions. These amino acid substitutions I analyzed are less than ten amino acid changes. Neurotransmitter receptors seems to have been generated by domain changes in some receptors and diversified by a small number of amino acid changes, and some of domains may have inherited from very old ancestors.

Fig. 3.24 (next page): Composite trees of all neurotransmitter receptors I analyzed.


Bacteria/
Eukaryotes
Plants/
Animals
Porifera/
Eumetazoa
Adenylate Cyclase
Bacteria
Soluble Guanylyl Cyclase
Soluble-GC
Guanylyl Cyclase-A, B
Natriuretic Peptide Receptor-A, B
Natriuretic Peptide Receptor-C
Guanylyl Cyclase-C
Guanylyl Cyclase-D, E, F, G
Eumetazoa
Glutamate Receptor
Taste Receptor
(T1R1, T1R2, T1R3)
Pheromone Receptor,
Odorant Receptor
Ca2+-Sensing Receptor
2.3
Eumetazoa
Animals
Bride of Sevenless Protein
GABA Receptor 2
GABA Receptor 1
Glutamate GABA-like Receptor
Porifera
Glutamate Receptor
(AMPA, Kainate)
Kainate Binding Protein
Glutamate Receptor (δ)
Glutamate Receptor (NMDA)
1.2
Eumetazoa
Glutamate Receptor-like Protein
Plants
Amino Acid BindingProtein
Bacteria
Amino Acid BindingProtein
GABA Receptor
Glycine Receptor
Histamine Receptor
1.1
Eumetazoa
5-HT Receptor
Acetylcholine Receptor
Acetylcholine Receptor
P2Y Receptor, Peptide Receptor (Thrombin),
Leukotriene Receptor, UDP-glucose Receptor,
Platelet Activating Factor Receptor,
Sphingosylphosphorylcholine Receptor
Peptide Receptor
(Chemokine, Angiotensin, Bradykinin)
Peptide Receptor
(Chemotactic Peptide, N-formyl Peptide)
Peptide Receptor
(Melatonin-concentrating Hormone, Opioid, Somatostatin)
Peptide Receptor (Adrenomedulin, Galanin)
Peptide Receptor
(Endothelin, Bombesin, Cholecystokinin,
Neuropeptide FF, Orexin, Tachykinin, Neuropeptide Y)
Peptide Receptor
(Thyrotropin-releasing Hormone, Neurotensin,
Growth Hormone Secretagogue, Motilin)
Peptide Receptor
(Glycoprotein Hormone)
Peptide Receptor (Relaxin)
Prostaglandin Receptor
Peptide Receptor
(Gonadotropin-releasing Hormone,
Urotensin II, Oxytocin, Vasopressin)
Peptide Receptor (Prokineticin)
Cannabinoid Receptor,
Lysophocphatidic Acid Receptor,
Melanocortin Receptor,
Adenosine Receptor
Acetylcholine Receptor,
Histamine Receptor,
Amine Receptor (Adrenaline, Dopamine, 5-HT, Octopamine)
Melatonin Receptor,
Olfactory Receptor,
Opsin Receptor
2.1
Eumetazoa
P2X Receptor
1.4
Vertebrates
Invertebrates/
Vertebrates
:Pkinase
:Guanylate_cyc
:7tm_3
:Sushi
:Lig_chan
:SBP_bac_3
:ANF_receptor
:Neur_chan_LBD
:Neur_chan_memb
:7tm_1
:LPR
:LPRNT
:Ldl_recept_a
:P2X_receptor

3.4.2 Evolution of Neurotransmission Systems

I investigated how neurotransmission systems have evolved. I first constructed phylogenetic trees of synthases, receptors and transporters of neurotransmitters. It seems that some synthases and some receptors, or some receptors and transporters have evolved together. Referring to topologies of each molecular phylogenetic tree (Fig. 3.25, 26, 27), it is possible to superimpose phylogenetic trees of S1 synthases and 2.1 receptors for histamine, dopamine and 5-HT, and also possible to impose phylogenetic trees of 2.1 receptors and T1 transporters for 5-HT, dopamine, adrenaline and noradrenaline (Fig. 3.28). It means that some of S1 synthases and 2.1 receptors might evolve together, and some of 2.1 receptors and T1 transporters might evolve together. If they evolved together, what had happened? I hypothesized about the co-evolution of S1 synthases and 2.1 receptors as follows. S1 synthases might evolve into the synthase for histamine and the synthase for dopamine and 5-HT at first. After that, amino acid substitutions might generate 2.1 receptors for histamine, dopamine and 5-HT. It is difficult to hypothesize the co-evolution of the 2.1 receptors and the T1 transporters, because I could not postulate which protein was generated at first. Although some of synthases and receptors or some of receptors and transporters might evolve together, most of proteins I analyzed seem to have evolved independently. Next, I focused on neurotransmitter receptors and did analysis of domain changes and amino acid substitutions which are essential for ligand binding specificity changes through evolution. Domain changes might be important for generating receptors in 1.2 and 2.3 receptors. A small number of amino acid substitutions might contribute to diversity of ligand binding specificities. Each protein may have evolved independently in neurotransmission system, and domain changes and a small number of amino acid

substitutions may have produced diverse neurotransmission systems.

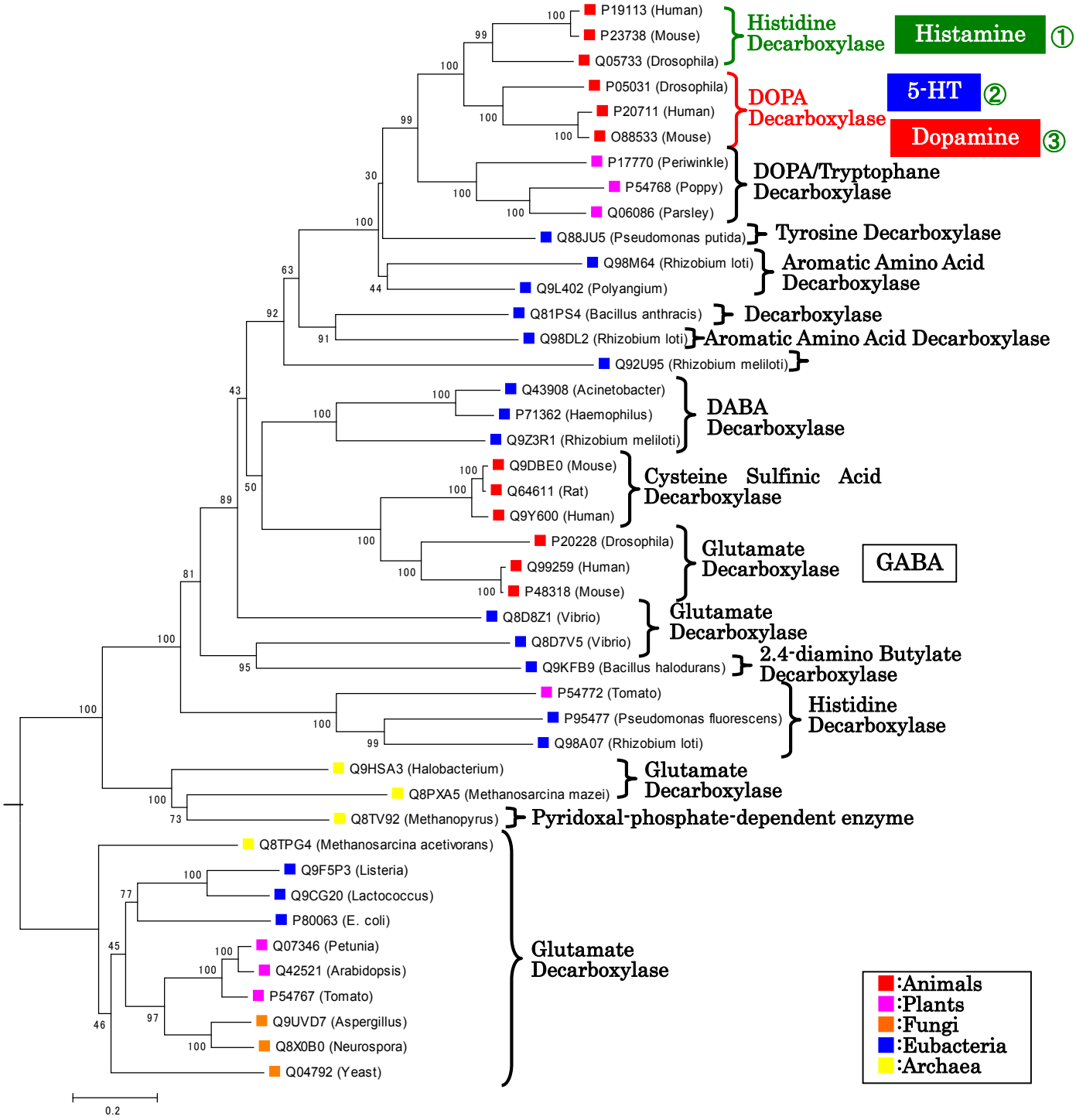


Fig. 3.25: Phylogenetic tree of S1 synthases

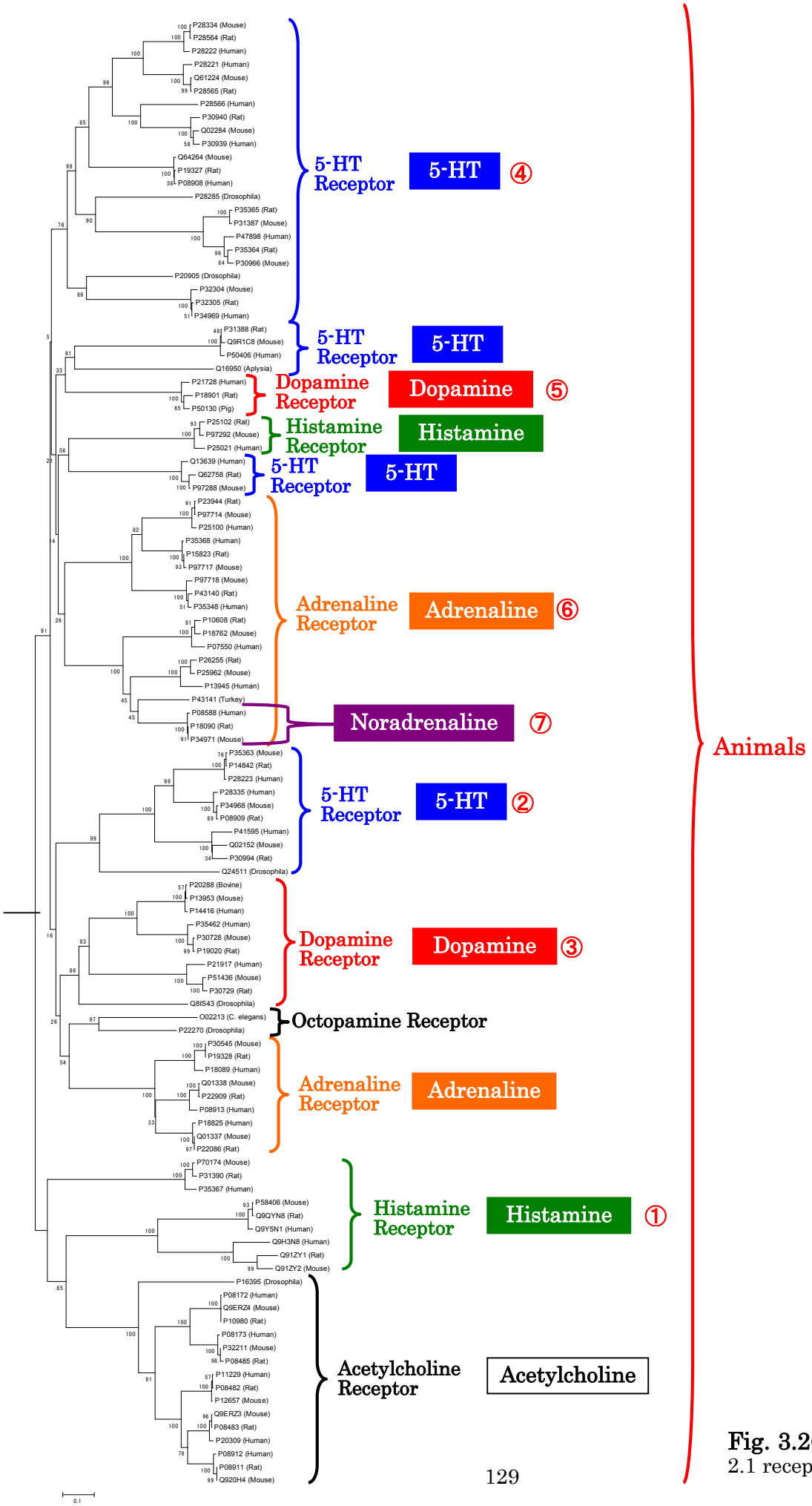
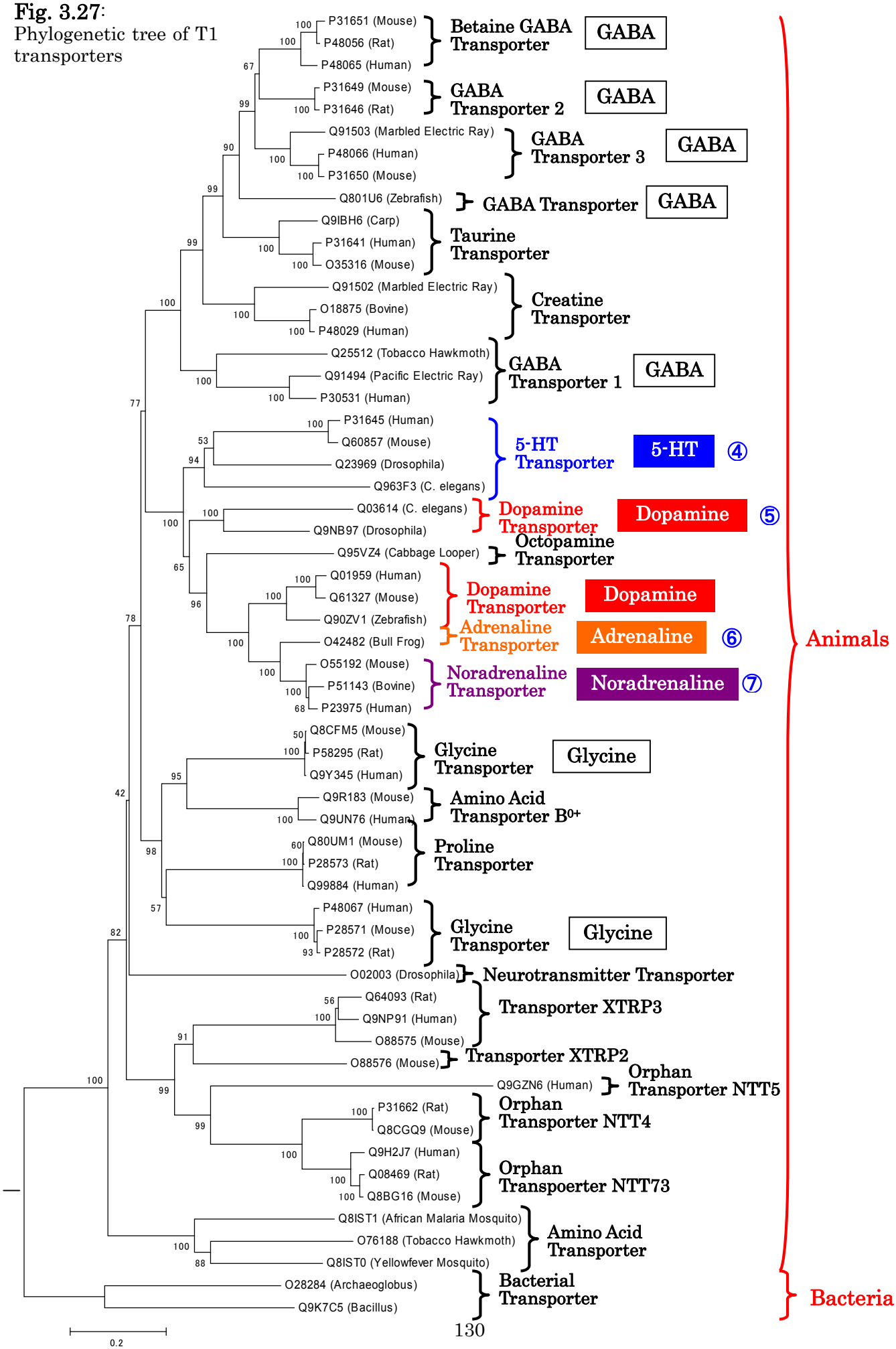


Fig. 3.26: Phylogenetic tree of 2.1 receptors

Fig. 3.27:
Phylogenetic tree of T1 transporters



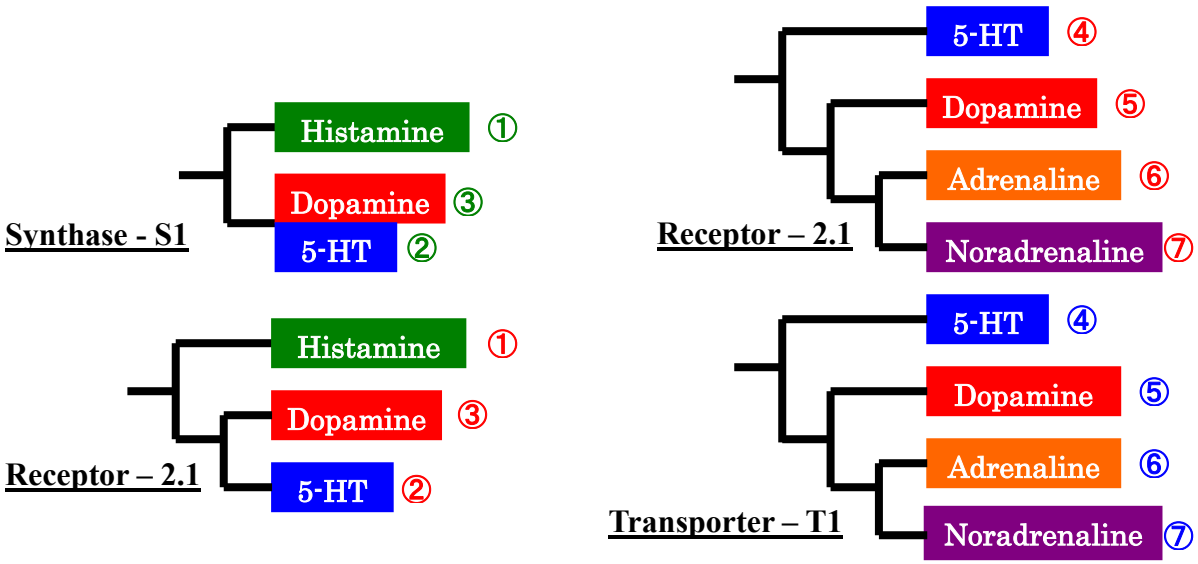


Fig. 3.28: Topologies of S1 synthases, 2.1 receptors and T1 transporters. Only proteins which are possible to co-evolve are shown. Each number corresponds to the numbers in Figs. 3.25-27.

4. Phylogenetic Analysis of Ion Channels

4.1 Introduction

Voltage-gated ion channels are important for generating action potentials in neurons. The evolution of voltage-gated potassium, calcium and sodium ion channels is thought to be like building blocks. Voltage-gated potassium channel is composed of four subunits, and each subunit has six hydrophobic segments, S1-S6 (Tempel et al. 1987). Four repeats of the six hydrophobic segments (S1-S6) are in one molecule of voltage-gated calcium and sodium channels which is a monomer channel (Noda et al. 1984; Tanabe et al. 1987). Voltage-gated calcium and sodium channels are hypothesized to have evolved by gene duplications from the ancestral protein that was similar to a potassium channel (Anderson and Greenberg 2001). For the purpose to infer the evolution of the voltage-gated channel family, I first conuducted the phylogenetic analysis of the potassium channels, the simplest channels among voltage-gated potassium, calcium and sodium channels.

The 12 sub-families in voltage-gated potassium channel (K_v) family including the 38 sub sub-families are defined by IUPHAR (Catterall et al. 2002). I did phylogenetic analysis of the four sub-families which were identified by early studies (Butler et al. 1989). The four sub-families are Kv1, Kv2, Kv3 and Kv4, and there are many experimental data of the channels in the four sub-families. By merging results of phylogenetic analysis and experimental data, I predicted amino acid substitutions which are important for changing the physiological features through the evolution of voltage-gated potassium channels. I especially focused on N-type inactivation of the channels which is one of major physiological features, and hypothesized how this

feature has been changed by amino acid substitutions through evolution.

4.2 Materials and Methods

4.2.1 Amino Acid Sequence Retrieval

I retrieved the amino acid sequences of voltage-gated potassium channels from SWISS-PROT and TrEMBL (Boeckmann et al. 2003).

4.2.2 Database Construction

Amino acid sequence data of voltage-gated potassium channels were integrated into a web-based SQL database (PostgreSQL) and I named it "Piment" (http://neuron.genes.nig.ac.jp/piment/), after the overall molecular shape of the voltage-gated potassium channel. "Piment" is a database including not only voltage-gated potassium channels but also other ion channels. Categories in the database are based on the bibliographies (Catterall et al. 2002; Conley 1996a; Conley 1996b; Conley and Brammer 1999). This database has five tables, and each table has relationships to other tables (lines between tables in Fig. 4.1). To entry data, administrator, namely myself, of Piment database inputs only two data; accession number and category data which are defined by myself. Categories in the database are mainly based on the homologous groups. Each protein information is linked to amino acid database on GenomeNet (http://www.genome.ad.jp) (① in Fig. 4.2). Amino acid sequence data of each protein is also linked to amino acid database on GenomeNet (② in Fig. 4.2). Users can easily get amino acid sequence data of proteins in Piment. After checking boxes of proteins of which you want to get sequences (③) and clicking get the sequences button (④), users can get sequences of checked proteins. The sequences can store in virtual basket of piment web site (⑤), and users can convert to some sequence formats, or do multiple alignments and construct phylogenetic tree by ClustalW on

GenomeNet or DDBJ.

T_INDEX

Field Name	ID †	DB*	OS	FRAG	PFAM	PDB	DE
Data Type	VARCHAR(20)	VARCHAR(3)	TEXT	INT2	TEXT	TEXT	TEXT

T_AC

Field Name	NUM †	AC*	ID*
Data Type	VARCHAR(20)	VARCHAR(20)	VARCHAR(20)

T_PROT

Field Name	AC †	CAT_ID*	PMID
Data Type	VARCHAR(20)	VARCHAR(39)	TEXT

T_CAT

Field Name	CAT_ID †	CAT_NAME#*	CAT_DE
Data Type	VARCHAR(39)	VARCHAR(50)	VARCHAR(254)

T_PUBMED

名称	PMID †	TITLE*	AUTH*	JN*
型	VARCHAR(10)	VARCHAR(254)	VARCHAR(254)	VARCHAR(254)

†: Primary Key, #: Unique, *: Not Null, Blue Characters: INDEX

Fig. 4.1: Tables in Macaroni database. Abbrebiations of the field name are as follows. ID = ID number of each sequence; DB = database where sequences are existed; OS = species name; FRAG = fragment data or not; PFAM = IDs of pfam domains which each protein has; PDB = pdb IDs, if tertiary structure of each protein is available; DE = description of each domain; NUM = non-redundant number for identification of each data; AC = accession number of each protein; CAT_ID = IDs of categories we defined; PMID = PubMed IDs of references in which functions of each protein describe; CAT_NAME = category name; CAT_DE = description of each category; TITLE = title of reference; AUTH = author of reference; JN = journal name of reference

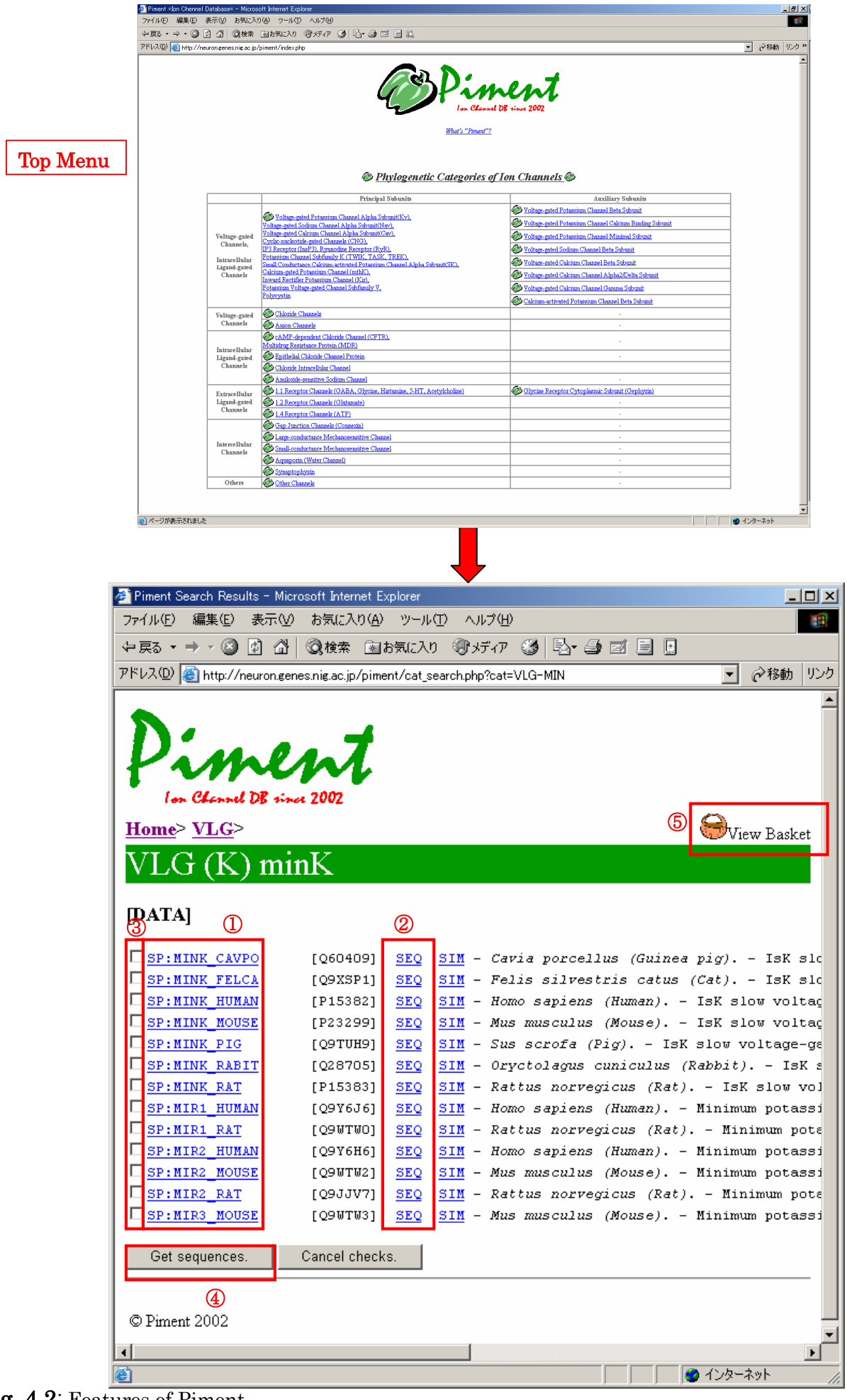


Fig. 4.2: Features of Piment

4.2.3 Domain Composition Analysis

I searched Pfam database (http://www.sanger.ac.uk/Software/Pfam/) (Bateman et al. 2002), the collection of protein families and domains, for functional domains of each protein.

4.2.4 Phylogenetic Analysis

The Piment database can collect FASTA format amino acid sequences of voltage-gated potassium channels which I select. I collected amino acid sequences by using the Piment database function and made multiple alignment by ClustalX (Thompson et al. 1997). Neighbor-joining tree (Saitou and Nei 1987) was constructed by using MEGA2 (Kumar et al. 2001). Numbers of amino acid substitutions were computed based on Poisson correction, because the evolutionary distances among voltage-gated potassium channels are short.

4.2.5 Analysis of Domain Changes and Amino Acid Substitutions

I analyzed the relationship between amino acid sequences and physiological functions. Physiological functions may have changed by several amino acid substitutions through evolution. I predicted how the domain compositions had changed through evolution referring to the molecular phylogenetic tree I constructed. I also focused on the substitutions of a small number of amino acids. I predicted the important amino acid substitutions for changing N-type inactivation, one of the major physiological functions, of voltage-gated potassium channels based on previous studies and the phylogenetic tree of the channels I constructed.

4.2.6 Analysis of N-terminal Amino Acid Chemical Features

Hydrophobicity and positive charge of the first 20 N-terminal amino acids are important for generating N-type inactivation. I standardized hydrophobicitcy value (Nozaki and Tanford 1971) and isoelectric point of each amino acid and calculated the average of each value in N-terminal amino acids. I did principal component analysis by using the average standardized hydrophobicity values and isoelectric points.

4.3 Results and Discussion

4.3.1 Inactivation of Voltage-gated Channels

Voltage-gated channels are important for generating action potential in neurons. Voltage-gated potassium channels of Kv1-4 are composed of four α subunits. In Kv1 channel, β subunits form a complex with the four α subunit. Voltage-gated channels have three states; resting, open and inactivation (Fig. 4.3). If membrane potential is changed, voltage-gated channel is activated and opens. After that voltage-gated channel is closed by inactivation. I focused on the inactivation, one of major physiological features of ion channels. There are two kinds of inactivation, N-type which is rapid inactivation and C-type which is slow inactivation (Rasmusson et al. 1998). The first 20 N-terminal amino acids in α subunit (Hoshi et al. 1990; Zagotta et al. 1990) and β-subunit (Rettig et al. 1994) act as the inactivation gate of the channel pore for generating N-type inactivation in Shaker-type voltage-gated potassium channels. The essential chemical features of the N-terminal amino acids to act as the inactivation gate are that the first approximately ten amino acids are hydrophobic, and the next ten amino acids are hydrophilic and positively charged (Murrell-Lagnado and Aldrich 1993a; Murrell-Lagnado and Aldrich 1993b).

4.3.2 Domain Changes

Fig. 4.4 shows the phylogenetic tree of voltage-gated potassium channel α subunits. Domain compositions are also shown in Fig. 4.4. All of the channels have Ion_trans domain which includes six hydrophobic segments (S1-S6) and K_tetra domain which is important to connect other α subunits for forming tetramer. The voltage-gated channels which are inactivated rapidly are colored red (bibliographies cited are listed in Table 4.1), and I suspected that these rapid inactivations are N-type inactivations. Domain compositions do not correspond to N-type inactivation channels, therefore the N-type inactivation was not generated by domain changes.

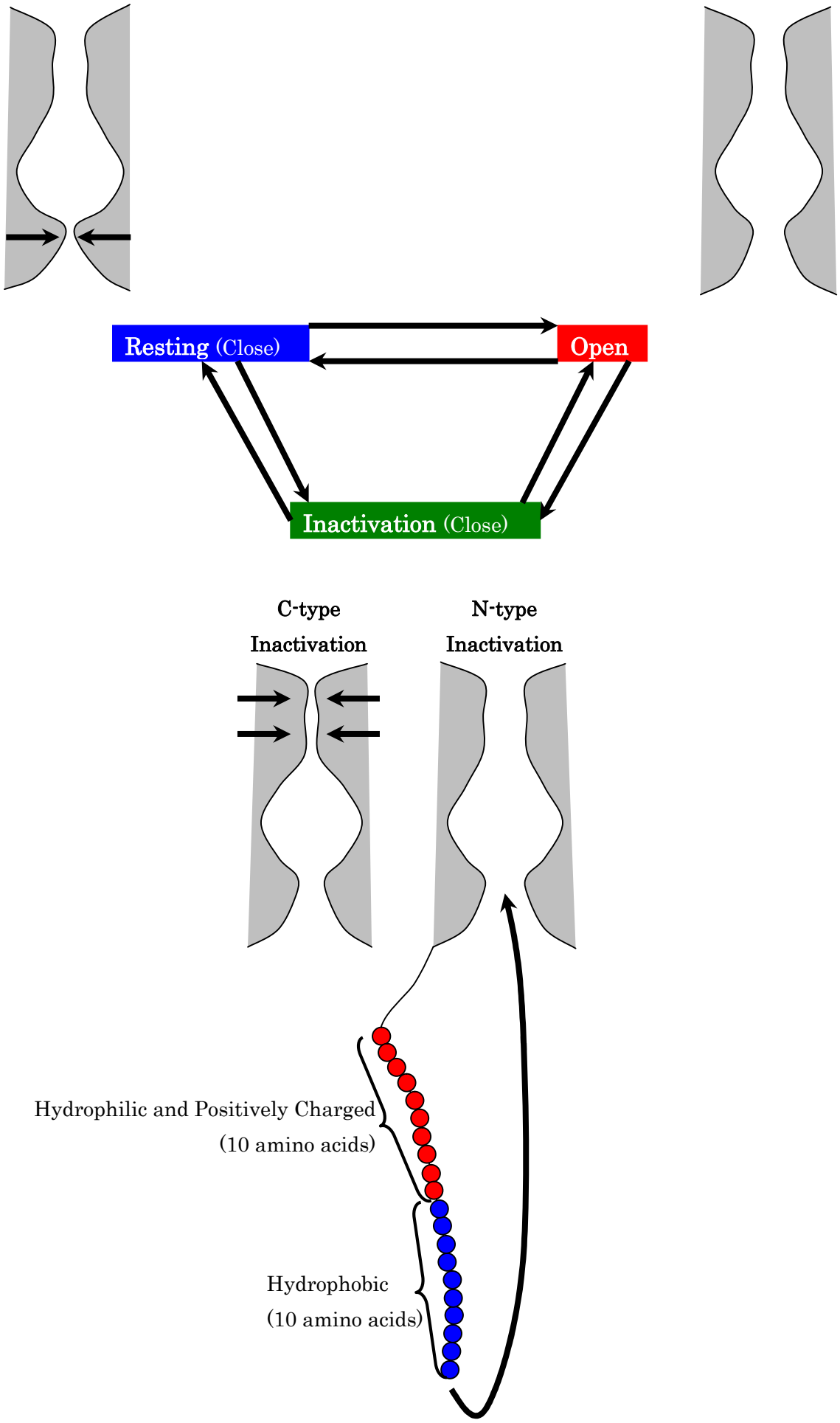


Fig. 4.3: Three states of ion channels.

Fig. 4.4: Phylogenetic tree of voltage-gated potassium channels of Kv1-4 subfamilies. Each colored square designates a Pfam domain. Channels whose names are written by red characters are inactivated rapidly, and the inactivations are suspected to be N-type inactivations. The red arrow on the lineage to Kv2.1 and 2.2 channels is predicted domain insertion.

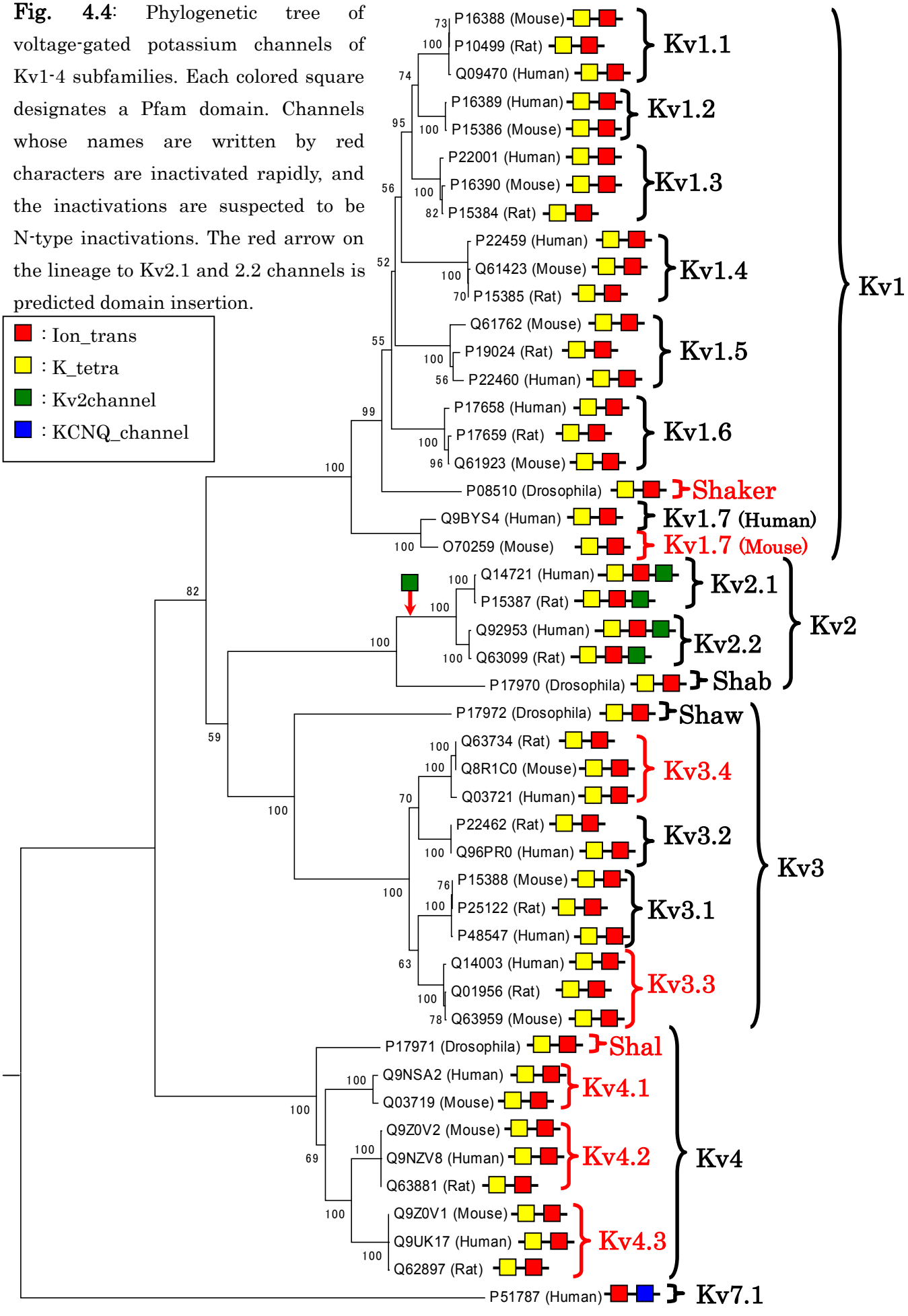


Table 4.1: References of electrophysiology data.

Channel	Reference
Shaker	Wei et al. 1990
Kv1.1	Akhtar et al. 2002; Christie et al. 1989; Grissmer et al. 1994; Stuhmer et al. 1989; Varshney et al. 2002; Zerr et al. 1998
Kv1.2	Akhtar et al. 2002; Grissmer et al. 1994; Koopmann et al. 2001; Paulmichl et al. 1991; Stuhmer et al. 1989; Varshney et al. 2002; Werkman et al. 1992
Kv1.3	Attali et al. 1992; Douglass et al. 1990; Grissmer et al. 1994; Stuhmer et al. 1989
Kv1.4	Jerng et al. 1999; Malin and Nerbonne 2002; Stuhmer et al. 1989; Tseng-Crank et al. 1990; Wymore et al. 1994
Kv1.5	Attali et al. 1993; Grissmer et al. 1994; Philipson et al. 1991
Kv1.6	Grupe et al. 1990; Kirsch et al. 1991
Kv1.7	Bardien-Kruger et al. 2002; Kalman et al. 1998
Shab	Pak et al. 1991b; Wei et al. 1990
Kv2.1	Frech et al. 1989; Koopmann et al. 2001; Malin and Nerbonne 2002; Martens et al. 2000; Murakoshi et al. 1997; Ottschytsch et al. 2002; Pak et al. 1991b; Zhu et al. 1999
Kv2.2	Hwang et al. 1992; Malin and Nerbonne 2002
Shaw	Wei et al. 1990
Kv3.1	Grissmer et al. 1992; Grissmer et al. 1994; Luneau et al. 1991b; Rettig et al. 1992; Weiser et al. 1994
Kv3.2	Luneau et al. 1991a; McCormack et al. 1990; Rettig et al. 1992; Vega-Saenz de Miera and Rudy 1992; Weiser et al. 1994
Kv3.3	Rae and Shepard 2000; Vega-Saenz de Miera and Rudy 1992; Weiser et al. 1994
Kv3.4	Shelton and Dart 1996; Weiser et al. 1994
Shal	Pak et al. 1991a; Wei et al. 1990
Kv4.1	Beck and Covarrubias 2001; Beck et al. 2002; Jerng et al. 1999; Pak et al. 1991a
Kv4.2	Martens et al. 2000; Shibata et al. 2000
Kv4.3	Beck and Covarrubias 2001; Beck et al. 2002; Tsaur et al. 1997

4.3.3 Amino Acid Changes Essential to N-type Inactivation of Voltage-gated Potassium Channels

As I described previous chapters, 20 N-terminal amino acids are important for generating amino acids. I calculated the average standardized hydrophobicity values of the first ten N-terminal amino acids (1-10 a.a.) and the average standardized hydrophobicity values and isoelectric points of the next ten N-terminal amino acids (11-20 a.a.) (Table 4.2, Fig. 4.3). The potassium ion channels which fulfilled conditions for occurring N-type inactivation are channels whose first ten amino acids are hydrophobic, and next ten amino acids are hydrophilic and positively charged. These channels are Shaker (Kv1), Kv1.5, Kv3.3 and Kv3.4, although Kv1.5 does not generate N-type inactivation (Table 4.3). Black circles in Table 4.3 are written in each column which is that the first ten amino acids are hydrophobic (the average standardized hydrophobicity values are more than 0), that the next ten amino acids are not hydrophobic (the average standardized hydrophobicity values are less than 0) or that the next ten amino acids are positively charged (the average standardized isoelectric points are more than 0). Channels which have three black circles have the specific 20 amino acids for generating N-type inactivation have the three black circles in Table 4.3, and they are Shaker, Kv.1.5, Kv3.3 and Kv3.4. These channels except for Kv1.5 are suddenly inactivated. Therefore, N-type inactivation is occurred in Shaker, Kv3.3 and Kv3.4. But channels which are inactivated suddenly (Kv1.7 (Mouse), Shal (Kv4), Kv4.1, Kv4.2 and Kv4.3) do not have the three black circles, that is, there are now specific amino acids for N-type inactivation. It suggests the N-type inactivations in Kv1.7 (Rat) and Kv4 sub-family channels are generated by different mechanisms from those in Shaker, Kv3.3 and Kv3.4.

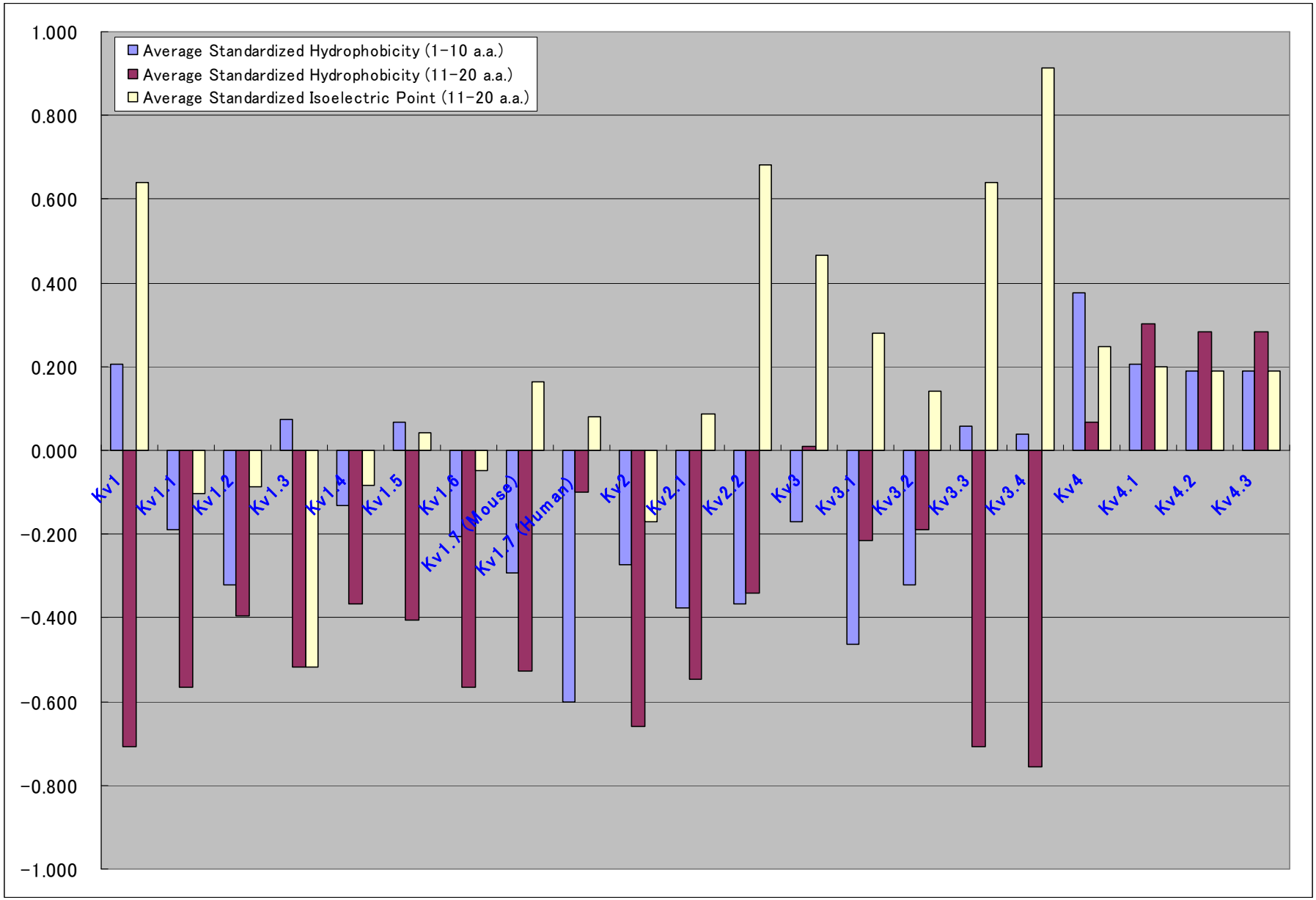


Fig. 4.5: Chemical features of the first 20 N-terminal amino acids in Kv1-4 channels. Average standardized hydrophobicity values in the first ten N-terminal amino acids (violet), average standardized hydrophobicity values in the next ten N-terminal amino acids (purple) and average standardized isoelectric points (light yellow) are shown.

Table 4.2: Standardized hydrophobicity values and standardized isoelectric points of each amino acid.

Amino Acids	Standerdized Hydrophobicity Values	Standerdized Isoelectric Points
A	−0.283	−0.030
C	−0.754	−0.556
D	−0.754	−1.858
E	−0.754	−1.603
F	1.603	−0.324
G	−0.754	−0.047
H	−0.283	0.870
I	0.943	−0.018
K	−0.754	2.087
L	0.943	−0.041
M	0.471	−0.177
N	−0.754	−0.364
P	−0.754	0.140
Q	−0.754	−0.228
R	−0.754	2.665
S	−0.754	−0.211
T	−0.377	0.061
V	0.660	−0.052
W	2.451	−0.092
Y	1.414	−0.222

Table 4.3: Chemical features in 20 N-terminal amino acids.

Channel Name	Species	Average Hydrophobicity		Average Isoelectric Point
		More than 0 (1–10 a.a.)	Less than 0 (11–20 a.a.)	More than 0 11–20 a.a.
Shaker (Kv1)	Drosophila	●	●	●
Kv1.1	Human	×	●	×
Kv1.2	Human	×	●	×
Kv1.3	Human	●	●	×
Kv1.4	Human	×	●	×
Kv1.5	Human	●	●	●
Kv1.6	Human	×	●	×
Kv1.7 (Mouse)	Mouse	×	●	●
Kv1.7 (Human)	Human	×	●	●
Shab (Kv2)	Drosophila	×	●	×
Kv2.1	Human	×	●	●
Kv2.2	Human	×	●	●
Shaw (Kv3)	Drosophila	×	×	●
Kv3.1	Human	×	●	●
Kv3.2	Human	×	●	●
Kv3.3	Human	●	●	●
Kv3.4	Human	●	●	●
Shal (Kv4)	Drosophila	●	×	●
Kv4.1	Human	●	×	●
Kv4.2	Human	●	×	●
Kv4.3	Human	●	×	●

I did principal component analysis by using three values; the average standardized hydrophobicity values of the first ten amino acids and the standardized hydrophobicity values and isoelectric points of the next ten amino acids. First, second and third principal components and percentages of inertia are shown in Table 4.4. Although inertia of the second principal component is not high, the second principal component is good for explaining the 20 specific amino acids which generate N-type inactivation. Because the second principal component is positively correlated with the hydrophobicity of the first ten amino acids and the isoelectric points of the next ten amino acids, and negatively correlated with the hydrophobicity of the next ten amino acids. It means the second principal component can explain the 20 specific amino acids including the first ten hydrophobic amino acids and the next ten hydrophilic and positively charged amino acids. The component scores of the second principal component are sorted by descending order and shown in Table 4.5. The three highest scores are the component scores of Kv3.4, Shaker and Kv3.3 which are inactivated by N-type inactivation. But the scores of Kv1.7 (Mouse), Shal, Kv4.1, Kv4.2 and Kv4.3 are not so high. As I hypothesized the previous paragraph, the mechanism of inactivation in Kv1.7 (Mouse), Shal, Kv4.1, Kv4.2 and Kv4.3 seems to be different from the ones in Shaker, Kv3.3 and Kv3.4.

Table 4.4: Eigen values, inertia and factor loading of principal component analyais.

		First	Second	Third
Eigen Values		1.288	1.042	0.670
Inertia (%)		0.429	0.347	0.223
Factor Loading	Hydrophobicity (1a.a.–10a.a.)	1.209	0.075	0.306
	Hydrophobicity (11a.a.–20a.a.)	1.009	−0.567	−0.269
	Isoelectric Points (11a.a.–20a.a.)	0.520	0.923	−0.190

Table 4.5: Component scores of second principal component.

Channel Name	Component Scores of Second Principal Component
Kv3.4	0.854
Shaker (Kv1)	0.611
Kv3.3	0.600
Kv2.2	0.415
Kv1.7 (Mouse)	0.078
Shaw (Kv3)	0.062
Kv2.1	0.016
Kv3.1	0.001
Kv1.5	-0.065
Kv1.6	-0.077
Shal (Kv4)	-0.115
Kv3.2	-0.120
Kv1.1	-0.122
Shab (Kv2)	-0.136
Kv1.4	-0.204
Kv1.2	-0.206
Kv1.7 (Human)	-0.238
Kv4.1	-0.291
Kv4.2	-0.291
Kv4.3	-0.291

How these specific amino acids for N-type inactivation were produced by amino acid substitutions? Fig. 4.6 shows the multiple alignment of Kv3 sub-family channels. Amino acid sequences of the N-terminal region in Kv3.3 and Kv3.4 channels are longer than those in Shaw (Kv3), Kv3,1 and Kv3.2, and these long sequences in Kv3.3 and Kv3.4 have the specific chemical features for generating N-type inactivation. Methionine at the translation initiation point of Shaw, Kv3.1 and Kv3.2 channels are conserved in all Kv3 channels. But the translation point of Kv3.3 and Kv3.4 channels is upstream from that of Shaw, Kv3.1 and Kv3.2. I hypothesized the scenario about the evolution of the Kv3 channels as follows. The N-terminal amino acids of Kv3.3 and Kv3.4 channels are similar, therefore these amino acids have been conserved from the ancestral protein of Kv3.3 and Kv3.4 channels. It means the ancestral protein of the Kv3.1, 3.2, 3.3 and 3.4 channels had the N-terminal amino acids like Kv3.3 and 3.4 channels. The initiation point in the ancestral protein of Kv3.1, 3.2, 3.3 and 3.4 channels was the one like Kv3.3 and 3.4, and the ancestral protein might be inactivated by N-type inactivation. Although the translation initiation point of the ancestral protein has been conserved in Kv3.3 and Kv3.4 channels, the point was changed on the lineages to Kv3.1 and Kv3.2 independently. And Kv3.1 and Kv3.2 lost the physiological feature of N-type inactivation.

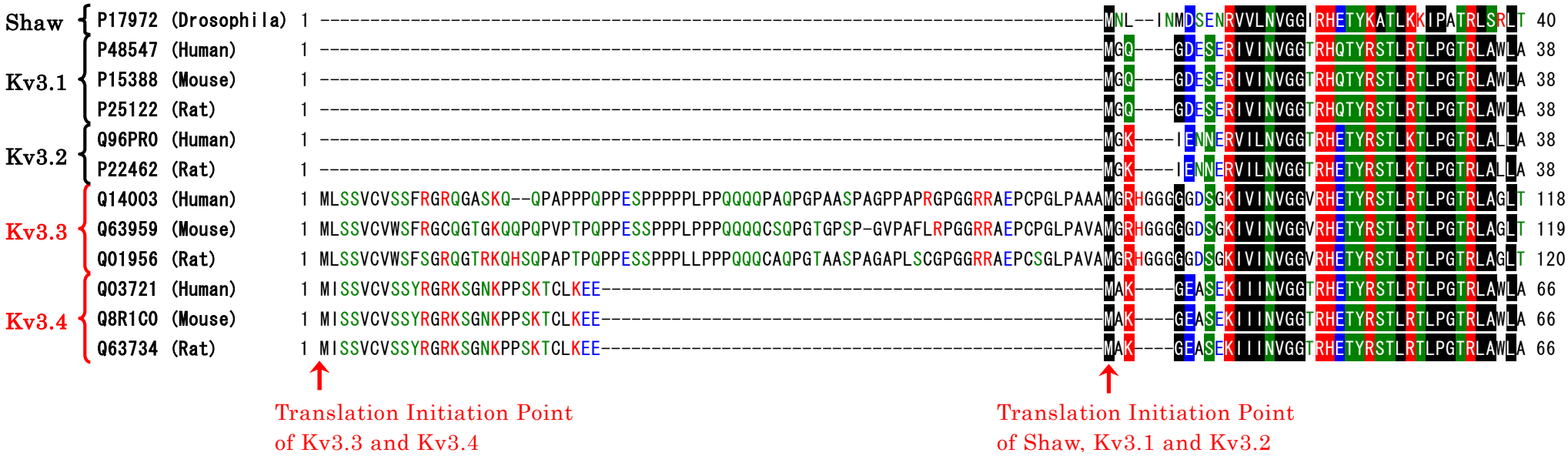


Fig. 4.6: Multiple alignment of N-terminal amino acids in Kv3 sub-family channels. Colors of each amino acid designate chemical feature as follows; red = basic amino acid (positively charged), blue = acidic amino acid (negatively charged), green = uncharged polar amino acid, black nonpolar amino acid. Channels whose names are written by red characters are N-type inactivation channels.

4.3.4 Conclusion

I focused on N-type inactivation, one of major features of ion channels. Domain changes might not influence the divergence of inactivation. A small number of amino acid substitutions might produce the N-type inactivation. Some of N-type inactivation channels which are Shaker (Kv1), Kv3.3 and Kv3.4 have the specific N-terminal amino acids, but other N-type inactivation channels which are Kv1.7 (Mouse), Shal (Kv4), Kv4.1, Kv4.2 and Kv4.3 do not. Therefore, the mechanism of N-type inactivation in Shaker, Kv3.3 and Kv3.4 seems to be different from other N-type inactivation channels. Although the mechanisms of N-type inactivation in Shaker, Kv3.3 and Kv3.4 seem to be similar, the origins of N-type inactivation are different from Shaker, and Kv3.3 and 3.4 by thinking parsimoniously with referring to molecular phylogenetic tree in Fig. 4.4. Amino acid substitution for generating N-type inactivation in Shaker may have been occurred on the lineage to Shaker after the divergence between vertebrate and invertebrate. Amino acid substitution for generating N-type inactivation in Kv3.3 and 3.4 may have been occurred on the lineage to Kv3.1-3.4 channels after the divergence between vertebrate and invertebrate, and the physiological feature of N-type inactivation may have been lost on the lineages to Kv3.1 and 3.2 independently before the divergence between Rodentia and Primates by the substitution of one amino acid, methionine. If the mechanisms of N-type inactivation in all Kv4 sub-family channels, amino acid substitutions for generating N-type inactivation in Kv4 sub-family channels might occur before the divergence between vertebrate and invertebrate. The amino acid substitutions of Kv4 sub-family are older than those of Shaker, and Kv3.3 and 3.4 channels.

Voltage-gated calcium and sodium channels which are homologous to

voltage-gated potassium channel have four repeats of Ion_trans domain (Fig. 4.4). And the domain composition of Kv7.1 channel is different from that of Kv1-4 channels (Fig. 4.4). Domain changes might occur among these channels and change physiological features of the channels. But the physiological feature of N-type inactivation in Kv1-4 channels may not have been changed by domain changes, but by a small number of amino acid substitutions. Domain changes and amino acid substitutions may have generated the divergence of voltage-gated ion channels.

5. Overall Conclusion

–How have functional systems in organisms evolved? –

Organisms including us are using various kinds of functional systems for living. In this study, I predicted how functional systems had been evolved and how protein functions in each functional system had changed through molecular phylogenetic analyses at three levels (system, domain and amino acid levels).

I did phylogenetic analyses at system level to predict phylogeny of functional systems by superimpositions of phylogenetic trees constructed by as many homologous proteins in the functional systems as possible. Under the assumption that the divergences of proteins in functional systems correspond to the divergences of functional systems, superimpositions of phylogenetic trees of proteins in functional systems were done, and the deduced phylogenetic trees of the functional systems were inferred.

Phylogenetic analyses at domain level by constructing composite gene trees were done so as to predict protein function changes in functional systems. Some of proteins are composed of multiple functional domains, and each domain might be inserted or deleted through evolution. Insertions or deletions of functional domains have changed protein functions. I tried to predict how domain compositions had changed through evolution and how protein functions had diversified by domain insertions and deletions through molecular phylogenetic analyses.

I did phylogenetic analyses at amino acid level to predict protein function changes in functional systems. It is possible that protein functions are different even if domain compositions of proteins are the same. I tried to predict the essential amino acid

substitutions for changing protein functions through evolution referring to phylogenetic trees and amino acid sequences which cannot be detected by domain level analysis.

I applied the phylogenetic analyses at three levels to the two functional systems; electron transfer energy metabolism system (chapter 2) and neurotransmission system (chapter 3). The four electron transfer energy metabolism systems (photosynthesis, aerobic respiration, denitrification and sulfur respiration) share common characteristic features. Furthermore, homologous proteins exist among the four systems. They suggest that the four systems are evolutionarily related. I tried the three level predictions in the four energy metabolism systems to infer the phylogeny of the four systems and the protein function changes in the four systems. I constructed molecular phylogenetic trees by using amino acid sequences of functional domains. These trees and amino acid compositions of proteins suggest that domain insertions and deletions in the four systems made functions of electron transfer in proteins change. I tried to predict ligand binging specificities of catalytic proteins at the amino acid level, but experimental data about ligand binding functions are not enough for doing such predictions. Therefore, I did not predict protein function changes at amino acid level. Most of proteins in the four systems are not homologous each other. Only some important proteins for generating energy are homologous among the four systems. It means that most of proteins have evolved independently, and few of proteins are conserved among the four systems. I tried to superimpose phylogenetic trees of homologous proteins in the four energy metabolism systems to predict the phylogeny of the four energy metabolisms, and the phylogeny of aerobic respiration, denitrification and sulfur respiration was predicted.

I also did the three level predictions in chemical neurotransmission system. I

tried to predict how the chemical neurotransmission systems have been diversified by means of superimposing the phylogenetic trees of synthases, receptors and transporters. The phylogenetic trees of some synthases and some receptors are possible to superimpose, and the phylogenetic trees of some receptors and some transporters are also possible to superimpose. These proteins might evolve together. But most of other proteins seem to have evolved independently from this study. Therefore, the unit of evolution is not system in chemical neurotransmission systems. I also did phylogenetic analyses of receptors at the domain level and the amino acid level. I inferred domain composition changes which had changed protein functions from the domain level analyses. Some domain changes seem to be essential for generating some receptors. The essential amino acid substitutions for changing ligand specificities were also predicted from the amino acid level analysis.

Based on the phylogenetic analyses of energy metabolism and chemical neurotransmission system, I did phylogenetic analysis of voltage-gated potassium channels at two levels; domain level and amino acid level (chapter 4). I focused on inactivation, one of the major electrophysiological features of voltage-gated ion channels, and predicted how the diversification of the inactivation had occurred. There are two kinds of inactivation in voltage-gated potassium channels. One is N-type inactivation which is sudden inactivation immediately after activation, the other one is C-type inactivation which is slow inactivation. Reffering to phylogenetic trees and domain compositions of voltage-gated potassium channels, domain composition changes seem not to affect inactivation differences. Previous studies suggest that the specific chemical features of 20 N-terminal amino acids are important for generating N-type inactivation. Therefore, I investigated the specific chemical features of 20 N-terminal amino acid

sequences in voltage-gated potassium channels. The specific chemical features of the 20 amino acids which induce N-type inactivation are found in the three subtypes out of the 21 subtypes. A small number of amino acid substitutions might produce the three N-type inactivation subtypes. The specific chemical features are not found in 20 N-terminal amino acids in the other five subclasses which can generate N-type inactivation. The amino acid substitutions which produce the five subtypes may be different from those of the three subtypes which have the specific chemical features in the 20 N-terminal amino acids.

I did phylogenetic analyses at system, domain and amino acid levels. From system level analysis, most of proteins in functional systems seem to be evolved independently. Although functions have been conserved in energy metabolism systems and chemical neurotransmission system, most of proteins in these systems have not been conserved. Domain composition changes seem to be slower than a small number of amino acid substitutions referring to phylogenetic analyses in this study. The combinations of the slower domain composition changes and the faster amino acid substitutions may have changed protein functions and these combinations might produce divergence of protein functions. But the combinations might happen independent of evolution of functional systems. Functional systems may have conserved their functions by the combinations and been diversified also by the combinations.

6. References

※Numbers in brackets show chapters in which references are cited.

Abaibou H, Pommier J, Benoit S, Giordano G, Mandrand-Berthelot MA (1995) Expression and characterization of the *Escherichia coli fdo* locus and a possible physiological role for aerobic formate dehydrogenase. J Bacteriol 177:7141-7149 [2]

Aichi M, Takatani N, Omata T (2001) Role of NtcB in activation of nitrate assimilation genes in the cyanobacterium *Synechocystis* sp. strain PCC 6803. J Bacteriol 183:5840-5847 [2]

Akhtar S, Shamotienko O, Papakosta M, Ali F, Dolly JO (2002) Characteristics of brain Kv1 channels tailored to mimic native counterparts by tandem linkage of alpha subunits: implications for K+ channelopathies. J Biol Chem 277:16376-16382 [4]

Alexander, SPH and Peters, JA (2000) Receptor and Ion Channel Nomenclature Supplement. Trends Pharmacol Sci Suppl:1-120 [3]

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410 [3]

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402 [2, 3]

Anand R, Conroy WG, Schoepfer R, Whiting P, Lindstrom J (1991) Neuronal nicotinic acetylcholine receptors expressed in Xenopus oocytes have a pentameric quaternary structure. J Biol Chem 266:11192-11198 [3]

Anderson PA, Greenberg RM (2001) Phylogeny of ion channels: clues to structure and

function. Comp Biochem Physiol B Biochem Mol Biol 129:17-28 [4]

Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, Podowski RM, Naslund AK, Eriksson AS, Winkler HH, Kurland CG (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396:133-140 [2]

Aoyagi M, Arvai AS, Ghosh S, Stuehr DJ, Tainer JA, Getzoff ED (2001) Structures of tetrahydrobiopterin binding-site mutants of inducible nitric oxide synthase oxygenase dimer and implicated roles of Trp457. Biochemistry 40:12826-12832 [2]

Arakaki AK, Ceccarelli EA, Carrillo N (1997) Plant-type ferredoxin-NADP+ reductases: a basal structural framework and a multiplicity of functions. FASEB J 11:133-140 [2]

Armstrong N, Gouaux E (2000) Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. Neuron 28:165-181 [3]

Armstrong N, Sun Y, Chen GQ, Gouaux E (1998) Structure of a glutamate-receptor ligand-binding core in complex with kainate. Nature 395:913-917 [3]

Attali B, Lesage F, Ziliani P, Guillemare E, Honore E, Waldmann R, Hugnot JP, Mattei MG, Lazdunski M, Barhanin J (1993) Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K+ channel. J Biol Chem 268:24283-24289 [4]

Attali B, Romey G, Honore E, Schmid-Alliana A, Mattei MG, Lesage F, Ricard P, Barhanin J, Lazdunski M (1992) Cloning, functional expression, and regulation of two K+ channels in human T lymphocytes. J Biol Chem 267:8650-8657 [4]

Avila J, Perez MD, Brito N, Gonzalez C, Siverio JM (1995) Cloning and disruption of the

YNR1 gene encoding the nitrate reductase apoenzyme of the yeast *Hansenula polymorpha*. FEBS Lett 366:137-142 [2]

Ayalon G, Stern-Bach Y (2001) Functional assembly of AMPA and kainate receptors is mediated by several discrete protein-protein interactions. Neuron 31:103-113 [3]

Back E, Burkhart W, Moyer M, Privalle L, Rothstein S (1988) Isolation of cDNA clones coding for spinach nitrite reductase: complete sequence and nitrate induction. Mol Gen Genet 212:20-26 [2]

Back E, Dunne W, Schneiderbauer A, de Framond A, Rastogi R, Rothstein SJ (1991) Isolation of the spinach nitrite reductase gene promoter which confers nitrate inducibility on GUS gene expression in transgenic tobacco. Plant Mol Biol 17:9-18 [2]

Bardien-Kruger S, Wulff H, Arieff Z, Brink P, Chandy KG, Corfield V (2002) Characterisation of the human voltage-gated potassium channel gene, KCNA7, a candidate gene for inherited cardiac disorders, and its exclusion as cause of progressive familial heart block I (PFHBI). Eur J Hum Genet 10:36-43 [4]

Barnard EA, Skolnick P, Olsen RW, Mohler H, Sieghart W, Biggio G, Braestrup C, Bateson AN, Langer SZ (1998) International Union of Pharmacology. XV. Subtypes of gamma-aminobutyric acidA receptors: classification on the basis of subunit structure and receptor function. Pharmacol Rev 50:291-313 [3]

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL (2002) The Pfam protein families database. Nucleic Acids Res 30:276-280 [2, 3, 4]

Batie CJ, Ballou DP, Correll CC (1991) Phthalate dioxygenase reductase and related flavin-iron-sulfur containing electron transferases. In: Muller F (ed) Chemistry

and biochemistry of flavoenzymes. CRC Press, Florida, pp. 543-556 [2]

Beck EJ, Bowlby M, An WF, Rhodes KJ, Covarrubias M (2002) Remodelling inactivation gating of Kv4 channels by KChIP1, a small-molecular-weight calcium-binding protein. J Physiol 538 :691-706 [4]

Beck EJ, Covarrubias M (2001) Kv4 channels exhibit modulation of closed-state inactivation in inside-out patches. Biophys J 81:867-883 [4]

Berks BC, Ferguson SJ, Moir JW, Richardson DJ (1995) Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions. Biochim Biophys Acta 1232:97-173 [2]

Berry EA, Guergova-Kuras M, Huang LS, Crofts AR (2000) Structure and function of cytochrome bc complexes. Annu Rev Biochem 69:1005-1075 [1, 2]

Bianchi V, Reichard P, Eliasson R, Pontis E, Krook M, Jornvall H, Haggard-Ljungquist E (1993) *Escherichia coli* ferredoxin NADP+ reductase: activation of *E. coli* anaerobic ribonucleotide reduction, cloning of the gene (*fpr*), and overexpression of the protein. J Bacteriol 175:1590-1595 [2]

Blasco F, Iobbi C, Giordano G, Chippaux M, Bonnefoy V (1989) Nitrate reductase of *Escherichia coli*: completion of the nucleotide sequence of the *nar* operon and reassessment of the role of the alpha and beta subunits in iron binding and electron transfer. Mol Gen Genet 218:249-256 [2]

Blasco F, Iobbi C, Ratouchniak J, Bonnefoy V, Chippaux M (1990) Nitrate reductases of *Escherichia coli*: sequence of the second nitrate reductase and comparison with that encoded by the *narGHJI* operon. Mol Gen Genet 222:104-111 [2]

Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M (2003) The

SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31:365-370 [2, 3, 4]

Boess FG, Monsma FJ Jr, Bourson A, Zwingelstein C, Sleight AJ (1998) Residues in transmembrane regions III and VI contribute to the 5-ht6 receptor ligand binding site. Ann N Y Acad Sci 861:242-243 [3]

Boess FG, Steward LJ, Steele JA, Liu D, Reid J, Glencorse TA, Martin IL (1997) Analysis of the ligand binding site of the 5-HT3 receptor using site directed mutagenesis: importance of glutamate 106. Neuropharmacology 36:637-647 [3]

Boileau AJ, Evers AR, Davis AF, Czajkowski C (1999) Mapping the agonist binding site of the GABAA receptor: evidence for a beta-strand. J Neurosci 19:4847-4854 [3]

Boorman JP, Groot-Kormelink PJ, Sivilotti LG (2000) Stoichiometry of human recombinant neuronal nicotinic receptors containing the b3 subunit expressed in Xenopus oocytes. J Physiol 529:565-577 [3]

Bormann J (2000) The 'ABC' of GABA receptors. Trends Pharmacol Sci 21:16-19 [3]

Boyington JC, Gladyshev VN, Khangulov SV, Stadtman TC, Sun PD (1997) Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocysteine, and an Fe_4S_4 cluster. Science 275:1305-1308 [2]

Brady CA, Stanford IM, Ali I, Lin L, Williams JM, Dubin AE, Hope AG, Barnes NM (2001) Pharmacological comparison of human homomeric 5-HT_{3A} receptors versus heteromeric 5-$HT_{3A/3B}$ receptors. Neuropharmacology 41:282-284 [3]

Bray RC, Adams B, Smith AT, Bennett B, Bailey S (2000) Reversible dissociation of thiolate ligands from molybdenum in an enzyme of the dimethyl sulfoxide reductase family. Biochemistry 39:11258-11269 [2]

Bredt DS, Hwang PM, Glatt CE, Lowenstein C, Reed RR, Snyder SH (1991) Cloned and

expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase. Nature 351:714-718 [2]

Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der Oost J, Smit AB, Sixma TK (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. Nature 411:269-276 [3]

Breton J, Berks BC, Reilly A, Thomson AJ, Ferguson SJ, Richardson DJ (1994) Characterization of the paramagnetic iron-containing redox centres of *Thiosphaera pantotropha* periplasmic nitrate reductase. FEBS Lett 345:76-80 [2]

Bruhl A, Haverkamp T, Gisselmann G, Schwenn JD (1996) A cDNA clone from *Arabidopsis thaliana* encoding plastidic ferredoxin:sulfite reductase. Biochim Biophys Acta 1295:119-124 [2]

Butler CS, Charnock JM, Bennett B, Sears HJ, Reilly AJ, Ferguson SJ, Garner CD, Lowe DJ, Thomson AJ, Berks BC, Richardson DJ (1999) Models for molybdenum coordination during the catalytic cycle of periplasmic nitrate reductase from *Paracoccus denitrificans* derived from EPR and EXAFS spectroscopy. Biochemistry 38:9000-9012 [2]

Cammack R, Weiner JH (1990) Electron paramagnetic resonance spectroscopic characterization of dimethyl sulfoxide reductase of *Escherichia coli*. Biochemistry 29:8410-8416 [2]

Catterall WA, Chandy KG and Gutman GA (eds) (2002) The IUPHAR Compendium of Voltage-gated Ion Channels. IUPHAR Media, Leeds, UK [4]

Cavalli A, Fanelli F, Taddei C, De Benedetti PG, Cotecchia S (1996) Amino acids of the alpha1B-adrenergic receptor involved in agonist binding: differences in docking catecholamines to receptor subtypes. FEBS Lett 399:9-13 [3]

Chatterton JE, Awobuluyi M, Premkumar LS, Takahashi H, Talantova M, Shin Y, Cui J, Tu S, Sevarino KA, Nakanishi N, Tong G, Lipton SA, Zhang D (2002) Excitatory glycine receptors containing the NR3 family of NMDA receptor subunits. Nature 415:793-798 [3]

Chow W, Ragan I, Robinson BH (1991) Determination of the cDNA sequence for the human mitochondrial 75-kDa Fe-S protein of NADH-coenzyme Q reductase. Eur J Biochem 201: 547-550 [2]

Christie MJ, Adelman JP, Douglass J, North RA (1989) Expression of a cloned rat brain potassium channel in Xenopus oocytes. Science 244:221-224 [4]

Coley C, Woodward R, Johansson AM, Strange PG, Naylor LH (2000) Effect of multiple serine/alanine mutations in the transmembrane spanning region V of the D2 dopamine receptor on ligand binding. J Neurochem 74:358-366 [3]

Conley EC (1996a) THE ION CHANNEL FactsBook I Extracellular Ligand-Gated Channels. Academic Press, London, UK [4]

Conley EC (1996b) THE ION CHANNEL FactsBook II Intracellular Ligand-Gated Channels. Academic Press, London, UK [4]

Conley EC and Brammer WJ (1999) THE ION CHANNEL FactsBook IV Voltage-gated Channels. Academic Press, London, UK [4]

Cooper E, Couturier S, Ballivet M (1991) Pentameric structure and subunit stoichiometry of a neuronal nicotinic acetylcholine receptor. Nature 350:235-238 [3]

Correll CC, Batie CJ, Ballou DP, Ludwig ML (1992) Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]. Science 258:1604-1610 [2]

Corringer PJ, Le Novere N, Changeux JP (2000) Nicotinic receptors at the amino acid level. Annu Rev Pharmacol Toxicol 40:431-458 [3]

Cox BA, Henningsen RA, Spanoyannis A, Neve RL, Neve KA (1992) Contributions of conserved serine residues to the interactions of ligands with dopamine D2 receptors. J Neurochem 59:627-635 [3]

Cramm R, Siddiqui RA, Friedrich B (1994) Primary sequence and evidence for a physiological function of the flavohemoprotein of *Alcaligenes eutrophus*. J Biol Chem 269:7349-7354 [2]

Crane BR, Arvai AS, Gachhui R, Wu C, Ghosh DK, Getzoff ED, Stuehr DJ, Tainer JA (1997a) The structure of nitric oxide synthase oxygenase domain and inhibitor complexes. Science 278:425-431 [2]

Crane BR, Arvai AS, Ghosh DK, Wu C, Getzoff ED, Stuehr DJ, Tainer JA (1998) Structure of nitric oxide synthase oxygenase dimer with pterin and substrate. Science 279:2121-2126 [2]

Crane BR, Getzoff ED (1996) The relationship between structure and function for the sulfite reductases. Curr Opin Struct Biol 6:744-756 [2]

Crane BR, Rosenfeld RJ, Arvai AS, Ghosh DK, Ghosh S, Tainer JA, Stuehr DJ, Getzoff ED (1999) N-terminal domain swapping and metal ion binding in nitric oxide synthase dimerization. EMBO J 18:6271-6281 [2]

Crane BR, Siegel LM, Getzoff ED (1995) Sulfite reductase structure at 1.6 A: evolution and catalysis for reduction of inorganic anions. Science 270:59-67 [2]

Crane BR, Siegel LM, Getzoff ED (1997b) Structures of the siroheme- and Fe_4S_4-containing active center of sulfite reductase in different states of oxidation: heme activation via reduction-gated exogenous ligand exchange. Biochemistry

36:12101-12119 [2]

Crane BR, Siegel LM, Getzoff ED (1997c) Probing the catalytic mechanism of sulfite reductase by X-ray crystallography: structures of the *Escherichia coli* hemoprotein in complex with substrates, inhibitors, intermediates, and products. Biochemistry 36:12120-12137 [2]

Crawford NM, Smith M, Bellissimo D, Davis RW (1988) Sequence and nitrate regulation of the *Arabidopsis thaliana* mRNA encoding nitrate reductase, a metalloflavoprotein with three functional domains. Proc Natl Acad Sci USA 85:5006-5010 [2]

Cromer BA, Morton CJ, Parker MW (2002) Anxiety over $GABA_A$ receptor structure relieved by AChBP. Trends Biochem Sci 27:280-287 [3]

Czjzek M, Dos Santos JP, Pommier J, Giordano G, Mejean V, Haser R (1998) Crystal structure of oxidized trimethylamine N-oxide reductase from *Shewanella massilia* at 2.5 A resolution. J Mol Biol 284:435-447 [2]

Dahl C, Kredich NM, Deutzmann R, Truper HG (1993) Dissimilatory sulphite reductase from *Archaeoglobus fulgidus*: physico-chemical properties of the enzyme and cloning, sequencing and analysis of the reductase genes. J Gen Microbiol 139:1817-1828 [2]

Dehmel U, Engesser KH, Timmis KN, Dwyer DF (1995) Cloning, nucleotide sequence, and expression of the gene encoding a novel dioxygenase involved in metabolism of carboxydiphenyl ethers in Pseudomonas pseudoalcaligenes POB310. Arch Microbiol 163:35-41 [2]

Dias JM, Than ME, Humm A, Huber R, Bourenkov GP, Bartunik HD, Bursakov S, Calvete J, Caldeira J, Carneiro C, Moura JJ, Moura I, Romao MJ (1999) Crystal

structure of the first dissimilatory nitrate reductase at 1.9 A solved by MAD methods. Structure Fold Des 7:65-79 [1, 2]

Dos Santos JP, Iobbi-Nivol C, Couillault C, Giordano G, Mejean V (1998) Molecular analysis of the trimethylamine N-oxide (TMAO) reductase respiratory system from a *Shewanella* species. J Mol Biol 284:421-433 [2]

Douglass J, Osborne PB, Cai YC, Wilkinson M, Christie MJ, Adelman JP (1990) Characterization and functional expression of a rat genomic DNA clone encoding a lymphocyte potassium channel. J Immunol 144:4841-4850 [4]

Elgoyhen AB, Johnson DS, Boulter J, Vetter DE, Heinemann S (1994) Alpha 9: an acetylcholine receptor with novel pharmacological properties expressed in rat cochlear hair cells. Cell 79:705-715 [3]

Elgoyhen AB, Vetter DE, Katz E, Rothlin CV, Heinemann SF, Boulter J (2001) alpha10: a determinant of nicotinic cholinergic receptor function in mammalian vestibular and cochlear mechanosensory hair cells. Proc Natl Acad Sci U S A 98:3501-3506 [3]

Elz S, Kramer K, Pertz HH, Detert H, ter Laak AM, Kuhne R, Schunack W (2000) Histaprodifens: synthesis, pharmacological in vitro evaluation, and molecular modeling of a new class of highly active and selective histamine H_1-receptor agonists. J Med Chem 43:1071-1084 [3]

Ennion S, Hagan S, Evans RJ (2000) The role of positively charged amino acids in ATP recognition by human $P2X_1$ receptors. J Biol Chem 275:29361-29367 [3]

Eschenbrenner M, Coves J, Fontecave M (1995a) The flavin reductase activity of the flavoprotein component of sulfite reductase from *Escherichia coli*. A new model for the protein structure. J Biol Chem 270:20550-20555 [2]

Eschenbrenner M, Coves J, Fontecave M (1995b) NADPH-sulfite reductase flavoprotein from *Escherichia coli*: contribution to the flavin content and subunit interaction. FEBS Lett 374:82-84 [2]

Finel M (1998) Organization and evolution of structural elements within complex I. Biochim Biophys Acta 1364:112-121 [2]

Fischmann TO, Hruza A, Niu XD, Fossetta JD, Lunn CA, Dolphin E, Prongay AJ, Reichert P, Lundell DJ, Narula SK, Weber PC (1999) Structural characterization of nitric oxide synthase isoforms reveals striking active-site conservation. Nat Struct Biol 6:233-242 [2]

Frech GC, VanDongen AM, Schuster G, Brown AM, Joho RH (1989) A novel potassium channel with delayed rectifier properties isolated from rat brain by expression cloning. Nature 340:642-645 [4]

Friebe A, Koesling D (2003) Regulation of nitric oxide-sensitive guanylyl cyclase. Circ Res 93:96-105 [3]

Fukui H, Ohta K, Yamamoto D (1999) Ligand recognition and activation mechanism of histamine H1 receptor. Nippon Yakurigaku Zasshi 113:289-297 [3]

Gantz I, DelValle J, Wang LD, Tashiro T, Munzert G, Guo YJ, Konda Y, Yamada T (1992) Molecular basis for the interaction of histamine with the histamine H2 receptor. J Biol Chem 267:20840-20843 [3]

Garton SD, Temple CA, Dhawan IK, Barber MJ, Rajagopalan KV, Johnson MK (2000) Resonance Raman characterization of biotin sulfoxide reductase. Comparing oxomolybdenum enzymes in the ME_2SO reductase family. J Biol Chem 275:6798-6805 [2]

Gassner GT, Johnson DA, Liu HW, Ballou DP (1996) Kinetics of the reductive

half-reaction of the iron-sulfur flavoenzyme CDP-6-deoxy-L-threo-D-glycero-4-hexulose-3-dehydrase reductase. Biochemistry 35:7752-7761 [2]

Ghosh DK, Crane BR, Ghosh S, Wolan D, Gachhui R, Crooks C, Presta A, Tainer JA, Getzoff ED, Stuehr DJ (1999) Inducible nitric oxide synthase: role of the N-terminal beta-hairpin hook and pterin-binding segment in dimerization and tetrahydrobiopterin interaction. EMBO J 18:6260-6270 [2]

Gisselmann G, Klausmeier P, Schwenn JD (1993) The ferredoxin:sulphite reductase gene from *Synechococcus* PCC7942. Biochim Biophys Acta 1144:102-106 [2]

Gon S, Giudici-Orticoni MT, Mejean V, Iobbi-Nivol C (2001) Electron transfer and binding of the c-type cytochrome TorC to the trimethylamine N-oxide reductase in *Escherichia coli*. J Biol Chem 276:11545-11551 [2]

Gon S, Patte JC, Mejean V, Iobbi-Nivol C (2000) The *torYZ* (*yecK bisZ*) operon encodes a third respiratory trimethylamine N-oxide reductase in *Escherichia coli*. J Bacteriol 182:5779-5786 [2]

Gotti C, Hanke W, Maury K, Moretti M, Ballivet M, Clementi F, Bertrand D (1994) Pharmacology and biophysical properties of alpha 7 and alpha 7-alpha 8 alpha-bungarotoxin receptor subtypes immunopurified from the chick optic lobe. Eur J Neurosci 6:1281-1291 [3]

Grissmer S, Ghanshani S, Dethlefs B, McPherson JD, Wasmuth JJ, Gutman GA, Cahalan MD, Chandy KG (1992) The Shaw-related potassium channel gene, Kv3.1, on human chromosome 11, encodes the type l K+ channel in T cells. J Biol Chem 267:20971-20979 [4]

Grissmer S, Nguyen AN, Aiyar J, Hanson DC, Mather RJ, Gutman GA, Karmilowicz

MJ, Auperin DD, Chandy KG (1994) Pharmacological characterization of five cloned voltage-gated K+ channels, types Kv1.1, 1.2, 1.3, 1.5, and 3.1, stably expressed in mammalian cell lines. Mol Pharmacol 45:1227-1234 [4]

Groot-Kormelink PJ, Boorman JP, Sivilotti LG (2001) Formation of functional alpha3beta4alpha5 human neuronal nicotinic receptors in Xenopus oocytes: a reporter mutation approach. Br J Pharmacol 134:789-796 [3]

Groot-Kormelink PJ, Luyten WH, Colquhoun D, Sivilotti LG (1998) A reporter mutation approach shows incorporation of the "orphan" subunit beta3 into a functional nicotinic receptor. J Biol Chem 273:15317-15320 [3]

Grove J, Tanapongpipat S, Thomas G, Griffiths L, Crooke H, Cole J (1996) *Escherichia coli* K-12 genes essential for the synthesis of c-type cytochromes and a third nitrate reductase located in the periplasm. Mol Microbiol 19:467-481 [2]

Gruez A, Pignol D, Zeghouf M, Coves J, Fontecave M, Ferrer JL, Fontecilla-Camps JC (2000) Four crystal structures of the 60 kDa flavoprotein monomer of the sulfite reductase indicate a disordered flavodoxin-like module. J Mol Biol 299:199-212 [2]

Grupe A, Schroter KH, Ruppersberg JP, Stocker M, Drewes T, Beckh S, Pongs O (1990) Cloning and expression of a human voltage-gated potassium channel. A novel member of the RCK potassium channel family. EMBO J 9:1749-1756 [4]

Guigliarelli B, Asso M, More C, Augier V, Blasco F, Pommier J, Giordano G, Bertrand P (1992) EPR and redox characterization of iron-sulfur centers in nitrate reductases A and Z from *Escherichia coli*. Evidence for a high-potential and a low-potential class and their relevance in the electron-transfer mechanism. Eur J Biochem 207:61-68 [2]

Hahne K, Haucke V, Ramage L, Schatz G (1994) Incomplete arrest in the outer

membrane sorts NADH-cytochrome b5 reductase to two different submitochondrial compartments. Cell 79:829-839 [2]

Haines DC, Tomchick DR, Machius M, Peterson JA (2001) Pivotal role of water in the mechanism of P450BM-3. Biochemistry 40:13456-13465 [2]

Haniu M, McManus ME, Birkett DJ, Lee TD, Shively JE (1989) Structural and functional analysis of NADPH-cytochrome P-450 reductase from human liver: complete sequence of human enzyme and NADPH-binding sites. Biochemistry 28:8639-8645 [2]

Hanna MC, Davies PA, Hales TG, Kirkness EF (2000) Evidence for expression of heteromeric serotonin 5-HT_3 receptors in rodents. J Neurochem 75:240-247 [3]

Hansen J, Cherest H, Kielland-Brandt MC (1994) Two divergent *MET10* genes, one from *Saccharomyces cerevisiae* and one from *Saccharomyces carlsbergensis*, encode the alpha subunit of sulfite reductase and specify potential binding sites for FAD and NADPH. J Bacteriol 176:6050-6058 [2]

Harayama S, Rekik M, Bairoch A, Neidle EL, Ornston LN (1991) Potential DNA slippage structures acquired during evolutionary divergence of *Acinetobacter calcoaceticus* chromosomal *benABC* and *Pseudomonas putida* TOL pWW0 plasmid *xylXYZ*, genes encoding benzoate dioxygenases. J Bacteriol 173:7540-7548 [2]

Hauska G, Nitschke W, Herrmann RG (1988) Amino acid identities in the three redox center-carrying polypeptides of cytochrome bc_1/b_6f complexes. J Bioenerg Biomembr 20:211-228 [1, 2]

Hazzard JT, Govindaraj S, Poulos TL, Tollin G (1997) Electron transfer between the FMN and heme domains of cytochrome P450BM-3. Effects of substrate and CO. J Biol Chem 272:7922-7926 [2]

Heffron K, Leger C, Rothery RA, Weiner JH, Armstrong FA (2001) Determination of an optimal potential window for catalysis by *E. coli* dimethyl sulfoxide reductase and hypothesis on the role of Mo(V) in the reaction pathway. Biochemistry 40:3117-3126 [2]

Heinzinger NK, Fujimoto SY, Clark MA, Moreno MS, Barrett EL (1995) Sequence analysis of the *phs* operon in *Salmonella typhimurium* and the contribution of thiosulfate reduction to anaerobic energy metabolism. J Bacteriol 177:2813-2820 [2]

Hipp WM, Pott AS, Thum-Schmitz N, Faath I, Dahl C, Truper HG (1997) Towards the phylogeny of APS reductases and sirohaem sulfite reductases in sulfate-reducing and sulfur-oxidizing prokaryotes. Microbiology 143:2891-2902 [2]

Hirai H, Kirsch J, Laube B, Betz H, Kuhse J (1996) The glycine binding site of the N-methyl-D-aspartate receptor subunit NR1: identification of novel determinants of co-agonist potentiation in the extracellular M3-M4 loop region. Proc Natl Acad Sci U S A 93:6031-6036 [3]

Hogg RC, Raggenbass M, Bertrand D (2003) Nicotinic acetylcholine receptors: from structure to brain function. Rev Physiol Biochem Pharmacol 147:1-46 [3]

Hoshi T, Zagotta WN, Aldrich RW (1990) Biophysical and molecular mechanisms of Shaker potassium channel inactivation. Science 250:533-538 [4]

Hsiao CD, Sun YJ, Rose J, Wang BC (1996) The crystal structure of glutamine-binding protein from Escherichia coli. J Mol Biol 262:225-242 [3]

Huang CJ, Barrett EL (1991) Sequence analysis and expression of the *Salmonella typhimurium asr* operon encoding production of hydrogen sulfide from sulfite. J Bacteriol 173:1544-1553 [2]

Humphrey PP, Barnard EA (1998) International Union of Pharmacology. XIX. The IUPHAR receptor code: a proposal for an alphanumeric classification system. Pharmacol Rev 50:271-277 [3]

Hwa J, Perez DM (1996) The unique nature of the serine interactions for alpha 1-adrenergic receptor agonist binding and activation. J Biol Chem 271:6322-6327 [3]

Hwang PM, Glatt CE, Bredt DS, Yellen G, Snyder SH (1992) A novel K+ channel with unique localizations in mammalian brain: molecular cloning and characterization. Neuron 8:473-481 [4]

Ingelman M, Bianchi V, Eklund H (1997) The three-dimensional structure of flavodoxin reductase from *Escherichia coli* at 1.7 A resolution. J Mol Biol 268:147-157 [2]

Isas JM, Burgess BK (1994) Purification and characterization of a NADP+/NADPH-specific flavoprotein that is overexpressed in FdI^- strains of *Azotobacter vinelandii*. J Biol Chem 269:19404-19409 [2]

Iwaasa H, Takagi T, Shikama K (1992) Amino acid sequence of yeast hemoglobin. A two-domain structure. J Mol Biol 227:948-954 [2]

Jansen T, Reilander H, Steppuhn J, Herrmann RG (1988) Analysis of cDNA clones encoding the entire precursor-polypeptide for ferredoxin:$NADP^+$ oxidoreductase from spinach. Curr Genet. 13:517-522 [2]

Janssens SP, Shimouchi A, Quertermous T, Bloch DB, Bloch KD (1992) Cloning and expression of a cDNA encoding human endothelium-derived relaxing factor/nitric oxide synthase. J Biol Chem 267:14519-14522 [2]

Jayaraman PS, Peakman TC, Busby SJ, Quincey RV, Cole JA (1987) Location and sequence of the promoter of the gene for the NADH-dependent nitrite reductase of

Escherichia coli and its regulation by oxygen, the Fnr protein and nitrite. J Mol Biol 196:781-788 [2]

Jentsch TJ, Stein V, Weinreich F, Zdebik AA (2002) Molecular structure and physiological function of chloride channels. Physiol Rev 82:503-568 [3]

Jerng HH, Shahidullah M, Covarrubias M (1999) Inactivation gating of Kv4 potassium channels: molecular interactions involving the inner vestibule of the pore. J Gen Physiol 113:641-660 [4]

Johnson JW, Ascher P (1987) Glycine potentiates the NMDA response in cultured mouse brain neurons. Nature 325 :529-531 [3]

Johnson DA, Gassner GT, Bandarian V, Ruzicka FJ, Ballou DP, Reed GH, Liu HW (1996) Kinetic characterization of an organic radical in the ascarylose biosynthetic pathway. Biochemistry 35:15846-15856 [2]

Johnstone IL, McCabe PC, Greaves P, Gurr SJ, Cole GE, Brow MA, Unkles SE, Clutterbuck AJ, Kinghorn JR, Innis MA (1990) Isolation and characterisation of the *crnA-niiA-niaD* gene cluster for nitrate assimilation in *Aspergillus nidulans*. Gene 90:181-192 [2]

Jones KA, Borowsky B, Tamm JA, Craig DA, Durkin MM, Dai M, Yao WJ, Johnson M, Gunwaldsen C, Huang LY, Tang C, Shen Q, Salon JA, Morse K, Laz T, Smith KE, Nagarathnam D, Noble SA, Branchek TA, Gerald C (1998) $GABA_B$ receptors function as a heteromeric assembly of the subunits $GABA_BR1$ and $GABA_BR2$. Nature 396:674-679 [3]

Jormakka M, Tornroth, S, Byrne B, Iwata S (2002) Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. Science 295:1863-1868 [2]

Kalman K, Nguyen A, Tseng-Crank J, Dukes ID, Chandy G, Hustad CM, Copeland NG, Jenkins NA, Mohrenweiser H, Brandriff B, Cahalan M, Gutman GA, Chandy KG (1998) Genomic organization, chromosomal localization, tissue distribution, and biophysical characterization of a novel mammalian Shaker-related voltage-gated potassium channel, Kv1.7. J Biol Chem 273:5851-5857 [4]

Kanehisa M, Goto S, Kawashima S, Nakaya A (2002) The KEGG databases at GenomeNet. Nucleic Acids Res 30:42-46 [2]

Karkhoff-Schweizer RR, Huber DP, Voordouw G (1995) Conservation of the genes for dissimilatory sulfite reductase from *Desulfovibrio vulgaris* and *Archaeoglobus fulgidus* allows their detection by PCR. Appl Environ Microbiol 61:290-296 [2]

Karlin A (2002) Emerging structure of the nicotinic acetylcholine receptors. Nat Rev Neurosci 3:102-114 [3]

Karlsson A, Beharry ZM, Matthew Eby D, Coulter ED, Neidle EL, Kurtz DM Jr., Eklund H, Ramaswamy S (2002) X-ray crystal structure of benzoate 1,2-dioxygenase reductase from *Acinetobacter* sp. strain ADP1. J Mol Biol 318:261-272 [2]

Karplus PA, Walsh KA, Herriott JR (1984) Amino acid sequence of spinach ferredoxin:$NADP^+$ oxidoreductase. Biochemistry. 23:6576-6583 [2]

Kaupmann K, Malitschek B, Schuler V, Heid J, Froestl W, Beck P, Mosbacher J, Bischoff S, Kulik A, Shigemoto R, Karschin A, Bettler B (1998) $GABA_B$-receptor subtypes assemble into functional heteromeric complexes. Nature396:683-687 [3]

Khakh BS, Burnstock G, Kennedy C, King BF, North RA, Seguela P, Voigt M, Humphrey PP (2001) International union of pharmacology. XXIV. Current status of the nomenclature and properties of P2X receptors and their subunits.

Pharmacol Rev 53:107-118 [3]

Kirsch GE, Drewe JA, Verma S, Brown AM, Joho RH (1991) Electrophysiological characterization of a new member of the RCK family of rat brain K+ channels. FEBS Lett 278:55-60 [4]

Kleckner NW, Dingledine R (1988) Requirement for glycine in activation of NMDA-receptors expressed in Xenopus oocytes. Science 241:835-837 [3]

Kniazeff J, Galvez T, Labesse G, Pin JP (2002) No ligand binding in the GB2 subunit of the GABA(B) receptor is required for activation and allosteric interaction between the subunits. J Neurosci 22:7352-7361 [3]

Koopmann R, Scholle A, Ludwig J, Leicher T, Zimmer T, Pongs O, Benndorf K (2001) Role of the S2 and S3 segment in determining the activation kinetics in Kv2.1 channels. J Membr Biol 182:49-59 [4]

Krafft T, Bokranz M, Klimmek O, Schroder I, Fahrenholz F, Kojro E, Kroger A (1992) Cloning and nucleotide sequence of the *psrA* gene of *Wolinella succinogenes* polysulphide reductase. Eur J Biochem 206:503-510 [1, 2]

Kramer H, Cagan RL, Zipursky SL (1991) Interaction of bride of sevenless membrane-bound ligand and the sevenless tyrosine-kinase receptor. Nature 352:207-212 [3]

Kuang D, Yao Y, Wang M, Pattabiraman N, Kotra LP, Hampson DR (2003) Molecular similarities in the ligand binding pockets of an odorant receptor and the metabotropic glutamate receptors. J Biol Chem 278:42551-42559 [3]

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244-1245 [2, 3, 4]

Kunishima N, Shimada Y, Tsuji Y, Sato T, Yamamoto M, Kumasaka T, Nakanishi S,

Jingami H, Morikawa K (2000) Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. Nature 407:971-977 [3]

Kuryatov A, Laube B, Betz H, Kuhse J (1994) Mutational analysis of the glycine-binding site of the NMDA receptor: structural similarity with bacterial amino acid-binding proteins. Neuron 12:1291-1300 [3]

Lafferty MA, Garrett RH (1974) Purification and properties of the *Neurospora crassa* assimilatory nitrite reductase. J Biol Chem 249:7555-7567 [2]

Lamas S, Marsden PA, Li GK, Tempst P, Michel T (1992) Endothelial nitric oxide synthase: molecular cloning and characterization of a distinct constitutive enzyme isoform. Proc Natl Acad Sci U S A 89:6348-6352 [2]

Larsen O, Lien T, Birkeland NK (1999) Dissimilatory sulfite reductase from *Archaeoglobus profundus* and *Desulfotomaculum thermocisternum*: phylogenetic and structural implications from gene sequences. Extremophiles 3:63-70 [2]

Laube B, Hirai H, Sturgess M, Betz H, Kuhse J (1997) Molecular determinants of agonist discrimination by NMDA receptor subunits: analysis of the glutamate binding site on the NR2B subunit. Neuron 18:493-503 [3]

Laube B, Kuhse J, Betz H (1998) Evidence for a tetrameric structure of recombinant NMDA receptors. J Neurosci 18:2954-2961 [3]

Le Novere N, Zoli M, Changeux JP (1996) Neuronal nicotinic receptor alpha 6 subunit mRNA is selectively concentrated in catecholaminergic nuclei of the rat brain. Eur J Neurosci 8:2428-2439 [3]

Leurs R, Smit MJ, Meeder R, Ter Laak AM, Timmerman H (1995) Lysine200 located in the fifth transmembrane domain of the histamine H1 receptor interacts with histamine but not with all H1 agonists. Biochem Biophys Res Commun

214:110-117 [3]

Li H, Poulos TL (1997) The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. Nat Struct Biol 4:140-146 [2]

Li H, Raman CS, Glaser CB, Blasko E, Young TA, Parkinson JF, Whitlow M, Poulos TL. (1999) Crystal structures of zinc-free and -bound heme domain of human inducible nitric-oxide synthase. Implications for dimer stability and comparison with endothelial nitric-oxide synthase. J Biol Chem 274:21276-21284 [2]

Li HK, Temple C, Rajagopalan KV, Schindelin H (2000) The 1.3 A crystal structure of *Rhodobacter sphaeroides* dimethyl sulfoxide reveals two distinct molybdenum coordination environments. J Am Chem Soc 122:7673-7680 [2]

Lin JT, Goldman BS, Stewart V (1993) Structures of genes *nasA* and *nasB*, encoding assimilatory nitrate and nitrite reductases in *Klebsiella pneumoniae* M5al. J Bacteriol 175:2370-2378 [2]

Lu G, Campbell WH, Schneider G, Lindqvist Y (1994) Crystal structure of the FAD-containing fragment of corn nitrate reductase at 2.5 A resolution: relationship to other flavoprotein reductases. Structure 2:809-821 [2]

Lu G, Lindqvist Y, Schneider G, Dwivedi U, Campbell W (1995) Structural studies on corn nitrate reductase: refined structure of the cytochrome b reductase fragment at 2.5 A, its ADP complex and an active-site mutant and modeling of the cytochrome b domain. J Mol Biol 248:931-948 [2]

Lu ZL, Saldanha JW, Hulme EC (2001) Transmembrane domains 4 and 7 of the M_1 muscarinic acetylcholine receptor are critical for ligand binding and the receptor activation switch. J Biol Chem 276:34098-34104 [3]

Lund J, Woodland MP, Dalton H (1985) Electron transfer reactions in the soluble methane monooxygenase of *Methylococcus capsulatus* (Bath). Eur J Biochem 147:297-305 [2]

Luneau C, Wiedmann R, Smith JS, Williams JB (1991a) Shaw-like rat brain potassium channel cDNA's with divergent 3' ends. FEBS Lett 288:163-167 [4]

Luneau CJ, Williams JB, Marshall J, Levitan ES, Oliva C, Smith JS, Antanavage J, Folander K, Stein RB, Swanson R, Kaczmarek LK, Buhrow SA (1991b) Alternative splicing contributes to K+ channel diversity in the mammalian central nervous system. Proc Natl Acad Sci U S A 88:3932-3936 [4]

Magalon A, Asso M, Guigliarelli B, Rothery RA, Bertrand P, Giordano G, Blasco F (1998) Molybdenum cofactor properties and [Fe-S] cluster coordination in *Escherichia coli* nitrate reductase A: investigation by site-directed mutagenesis of the conserved his-50 residue in the NarG subunit. Biochemistry 37:7363-7370 [2]

Magalon A, Rothery RA, Giordano G, Blasco F, Weiner JH (1997) Characterization by electron paramagnetic resonance of the role of the *Escherichia coli* nitrate reductase (NarGHI) iron-sulfur clusters in electron transfer to nitrate and identification of a semiquinone radical intermediate. J Bacteriol 179:5037-5045 [2]

Malin SA, Nerbonne JM (2002) Delayed rectifier K+ currents, IK, are encoded by Kv2 alpha-subunits and regulate tonic firing in mammalian sympathetic neurons. J Neurosci 22:10094-10105 [4]

Manivet P, Schneider B, Smith JC, Choi DS, Maroteaux L, Kellermann O, Launay JM (2002) The serotonin binding site of human and murine 5-HT2B receptors: molecular modeling and site-directed mutagenesis. J Biol Chem 277:17170-17178 [3]

Martens JR, Navarro-Polanco R, Coppock EA, Nishiyama A, Parshley L, Grobaski TD, Tamkun MM (2000) Differential targeting of Shaker-like potassium channels to lipid rafts. J Biol Chem 275:7443-7446 [4]

Martinez-Julvez M, Hermoso J, Hurley JK, Mayoral T, Sanz-Aparicio J, Tollin G, Gomez-Moreno C, Medina M (1998) Role of Arg100 and Arg264 from *Anabaena* PCC 7119 ferredoxin-$NADP^+$ reductase for optimal $NADP^+$ binding and electron transfer. Biochemistry 37:17680-17691 [2]

Martinez-Julvez M, Nogues I, Faro M, Hurley JK., Brodie TB, Mayoral T, Sanz-Aparicio J, Hermoso JA, Stankovich MT, Medina M, Tollin G, Gomez-Moreno C (2001) Role of a cluster of hydrophobic residues near the FAD cofactor in *Anabaena* PCC 7119 ferredoxin-$NADP^+$ reductase for optimal complex formation and electron transfer to ferredoxin. J Biol Chem 276:27498-27510 [2]

Mayat E, Petralia RS, Wang YX, Wenthold RJ (1995) Immunoprecipitation, immunoblotting, and immunocytochemistry studies suggest that glutamate receptor delta subunits form novel postsynaptic receptor complexes. J Neurosci 15:2533-2546 [3]

Mayoral T, Medina M, Sanz-Aparicio J, Gomez-Moreno C, Hermoso JA (2000) Structural basis of the catalytic role of Glu301 in *Anabaena* PCC 7119 ferredoxin-$NADP^+$ reductase revealed by x-ray crystallography. Proteins 38:60-69 [2]

McAlpine AS, McEwan AG, Bailey S (1998) The high resolution crystal structure of DMSO reductase in complex with DMSO. J Mol Biol 275:613-623 [2]

McCormack T, Vega-Saenz de Miera EC, Rudy B (1990) Molecular cloning of a member of a third class of Shaker-family K+ channel genes in mammals. Proc Natl Acad

Sci U S A 87:5227-5231 [4]

McMillan K, Adler M, Auld DS, Baldwin JJ, Blasko E, Browne LJ, Chelsky D, Davey, D, Dolle RE, Eagen KA, Erickson S, Feldman RI, Glaser CB, Mallari C, Morrissey MM, Ohlmeyer MH, Pan G, Parkinson JF, Phillips GB, Polokoff MA, Sigal NH, Vergona R, Whitlow M, Young TA, Devlin JJ (2000) Allosteric inhibitors of inducible nitric oxide synthase dimerization discovered via combinatorial chemistry. Proc Natl Acad Sci USA 97:1506-1511 [2]

McNicoll N, Gagnon J, Rondeau JJ, Ong H, De Lean A (1996) Localization by photoaffinity labeling of natriuretic peptide receptor-A binding domain. Biochemistry 35:12950-12956 [3]

Merchan F, Prieto R, Kindle KL, Llama MJ, Serra JL, Fernandez E (1995) Isolation, sequence and expression in *Escherichia coli* of the nitrite reductase gene from the filamentous, thermophilic cyanobacterium *Phormidium laminosum*. Plant Mol Biol 27:1037-1042 [2]

Mezler M, Muller T, Raming K (2001) Cloning and functional expression of $GABA_B$ receptors from *Drosophila*. Eur J Neurosci 13:477-486 [3]

Mialet J, Dahmoune Y, Lezoualc'h F, Berque-Bestel I, Eftekhari P, Hoebeke J, Sicsic S, Langlois M, Fischmeister R (2000) Exploration of the ligand binding site of the human 5-HT_4 receptor by site-directed mutagenesis and molecular modeling. Br J Pharmacol 130:527-538 [3]

Middleton RE, Cohen JB (1991) Mapping of the acetylcholine binding site of the nicotinic acetylcholine receptor: [3H]nicotine as an agonist photoaffinity label. Biochemistry 30:6987-6997 [3]

Mishina M, Takai T, Imoto K, Noda M, Takahashi T, Numa S, Methfessel C, Sakmann B

(1986) Molecular distinction between fetal and adult forms of muscle acetylcholine receptor. Nature 321:406-411[3]

Mogi T, Tsubaki M, Hori H, Miyoshi H, Nakamura H, Anraku Y (1998) Two Terminal Quinol Oxidase Families in *Escherichia coli*: Variations on Molecular Machinery for Dioxygen Reduction. J Biochem Mol Biol Biophys 2:79-110 [1, 2]

Morales R, Kachalova G, Vellieux F, Charon MH, Frey M (2000) Crystallographic studies of the interaction between the ferredoxin-$NADP^+$ reductase and ferredoxin from the cyanobacterium *Anabaena*: looking for the elusive ferredoxin molecule. Acta Crystallogr D Biol Crystallogr 56:1408-1412 [2]

Moreno-Vivian C, Cabello P, Martinez-Luque M, Blasco R, Castillo F (1999) Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. J Bacteriol 181:6573-6584 [1, 2]

Murakoshi H, Shi G, Scannevin RH, Trimmer JS (1997) Phosphorylation of the Kv2.1 K+ channel alters voltage-dependent activation. Mol Pharmacol 52:821-828 [4]

Murrell-Lagnado RD, Aldrich RW (1993a) Interactions of amino terminal domains of Shaker K channels with a pore blocking site studied with synthetic peptides. J Gen Physiol 102:949-975 [1, 4]

Murrell-Lagnado RD, Aldrich RW (1993b) Energetics of Shaker K channels block by inactivation peptides. J Gen Physiol 102:977-1003 [1, 4]

Nederkoorn PH, van Lenthe JH, van der Goot H, Donne-Op den Kelder GM, Timmerman H (1996) The agonistic binding site at the histamine H2 receptor. I. Theoretical investigations of histamine binding to an oligopeptide mimicking a part of the fifth transmembrane alpha-helix. J Comput Aided Mol Des 10:461-478 [3]

Newman BJ, Gray JC (1988) Characterization of a full-length cDNA clone for pea ferredoxin-NADP+ reductase. Plant Mol Biol 10:511-520 [2]

Noda M, Shimizu S, Tanabe T, Takai T, Kayano T, Ikeda T, Takahashi H, Nakayama H, Kanaoka Y, Minamino N, Kangawa K, Matsuo H, Raftery MA, Hirose T, Inayama S, Hayashida H, Miyata T, Numa S (1984) Primary structure of Electrophorus electricus sodium channel deduced from cDNA sequence. Nature 312:121-127 [4]

Nomura Y, Nakagawa M, Ogawa N, Harashima S, Oshima Y (1992) Genes in PHT Plasmid Encoding the Initial Degradation Pathway of Phthalate in *Pseudominas putida*. J Ferment Bioeng 74:333-344 [2]

Nozaki Y, Tanford C (1971) The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. J Biol Chem 246:2211-2217 [4]

North RA (2002) Molecular physiology of P2X receptors. Physiol Rev 82:1013-1067 [3]

Ogawa K, Akagawa E, Yamane K, Sun ZW, LaCelle M, Zuber P, Nakano MM (1995) The *nasB* operon and *nasA* gene are required for nitrate/nitrite assimilation in *Bacillus subtilis*. J Bacteriol 177:1409-1413 [2]

Ohnishi T (1998) Iron-sulfur clusters/semiquinones in complex I. Biochim Biophys Acta 1364:186-206 [2]

Ohta K, Hayashi H, Mizuguchi H, Kagamiyama H, Fujimoto K, Fukui H (1994) Site-directed mutagenesis of the histamine H1 receptor: roles of aspartic acid107, asparagine198 and threonine194. Biochem Biophys Res Commun 203:1096-1101 [3]

OOta S, Saitou N (1999) Phylogenetic relationship of muscle tissues deduced from superimposition of gene trees. Mol Biol Evol 16:856-867 [1]

Ostrowski J, Barber MJ, Rueger DC, Miller BE, Siegel LM, Kredich NM (1989a) Characterization of the flavoprotein moieties of NADPH-sulfite reductase from *Salmonella typhimurium* and *Escherichia coli*. Physicochemical and catalytic properties, amino acid sequence deduced from DNA sequence of *cysJ*, and comparison with NADPH-cytochrome P-450 reductase. J Biol Chem 264:15796-15808 [2]

Ostrowski J, Wu JY, Rueger DC, Miller BE, Siegel LM, Kredich NM (1989b) Characterization of the *cysJIH* regions of *Salmonella typhimurium* and *Escherichia coli* B. DNA sequences of *cysI* and *cysH* and a model for the siroheme-Fe_4S_4 active center of sulfite reductase hemoprotein based on amino acid homology with spinach nitrite reductase. J Biol Chem 264:15726-15737 [2]

Oswald RE, Changeux JP (1982) Crosslinking of alpha-bungarotoxin to the acetylcholine receptor from *Torpedo marmorata* by ultraviolet light irradiation. FEBS Lett 139:225-229 [3]

Ottschytsch N, Raes A, Van Hoorick D, Snyders DJ (2002) Obligatory heterotetramerization of three previously uncharacterized Kv channel alpha-subunits identified in the human genome. Proc Natl Acad Sci U S A 99:7986-7991 [4]

Ozols J, Carr SA, Strittmatter P (1984) Identification of the NH_2-terminal blocking group of NADH-cytochrome b_5 reductase as myristic acid and the complete amino acid sequence of the membrane-binding domain. J Biol Chem 259:13349-13354 [2]

Ozols J, Korza G, Heinemann FS, Hediger MA, Strittmatter P (1985) Complete amino acid sequence of steer liver microsomal NADH-cytochrome b_5 reductase. J Biol Chem 260:11953-11961 [2]

Page KM, Curtis CA, Jones PG, Hulme EC (1995) The functional role of the binding site aspartate in muscarinic acetylcholine receptors, probed by site-directed mutagenesis. Eur J Pharmacol 289:429-437 [3]

Pak MD, Baker K, Covarrubias M, Butler A, Ratcliffe A, Salkoff L (1991a) mShal, a subfamily of A-type K+ channel cloned from mammalian brain. Proc Natl Acad Sci U S A 88:4386-4390 [4]

Pak MD, Covarrubias M, Ratcliffe A, Salkoff L (1991b) A mouse brain homolog of the Drosophila Shab K+ channel with conserved delayed-rectifier properties. J Neurosci 11:869-880 [4]

Park J, Karplus K, Barrett C, Hughey R, Haussler D, Hubbard T, Chothia C (1998) Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. J Mol Biol 284:1201-1210 [2]

Paulmichl M, Nasmith P, Hellmiss R, Reed K, Boyle WA, Nerbonne JM, Peralta EG, Clapham DE (1991) Cloning and expression of a rat cardiac delayed rectifier potassium channel. Proc Natl Acad Sci U S A 88:7892-7895 [4]

Peakman T, Crouzet J, Mayaux JF, Busby S, Mohan S, Harborne N, Wootton J, Nicolson R, Cole J (1990) Nucleotide sequence, organisation and structural analysis of the products of genes in the *nirB-cysG* region of the *Escherichia coli* K-12 chromosome. Eur J Biochem 191:315-323 [2]

Peterson JA, Sevrioukova I, Truan G, Graham-Lorence SE (1997) P450BM-3; a tale of two domains--or is it three? Steroids 62:117-123 [2]

Philipson LH, Hice RE, Schaefer K, LaMendola J, Bell GI, Nelson DJ, Steiner DF (1991) Sequence and functional expression in Xenopus oocytes of a human insulinoma and islet potassium channel. Proc Natl Acad Sci U S A 88:53-57 [4]

Pierik AJ, Duyvis MG, van Helvoort JM, Wolbert RB, Hagen WR (1992) The third subunit of desulfoviridin-type dissimilatory sulfite reductases. Eur J Biochem 205:111-115 [2]

Pierson DE, Campbell A (1990) Cloning and nucleotide sequence of *bisC*, the structural gene for biotin sulfoxide reductase in *Escherichia coli*. J Bacteriol 172:2194-2198 [2]

Pieterse CM, van't Klooster J, van den Berg-Velthuis GC, Govers F (1995) *NiaA*, the structural nitrate reductase gene of *Phytophthora infestans*: isolation, characterization and expression analysis in *Aspergillus nidulans*. Curr Genet 27:359-366 [2]

Pietrini G, Carrera P, Borgese N (1988) Two transcripts encode rat cytochrome b_5 reductase. Proc Natl Acad Sci U S A 85:7246-7250 [2]

Pollock VV, Barber MJ (2001) Kinetic and mechanistic properties of biotin sulfoxide reductase. Biochemistry 40:1430-1440 [2]

Porter TD, Beck TW, Kasper CB (1990) NADPH-cytochrome P-450 oxidoreductase gene organization correlates with structural domains of the protein. Biochemistry 29:9814-9818 [2]

Porter TD, Kasper CB (1985) Coding nucleotide sequence of rat NADPH-cytochrome P-450 oxidoreductase cDNA and identification of flavin-binding domains. Proc Natl Acad Sci U S A 82:973-977 [2]

Powlowski J, Shingler V (1990) In vitro analysis of polypeptide requirements of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. J Bacteriol 172:6834-6840 [2]

Preis D, Weidner U, Conzen C, Azevedo JE, Nehls U, Rohlen D, van der Pas J,

Sackmann U, Schneider R, Werner S, Weiss H (1991) Primary structures of two subunits of NADH: ubiquinone reductase from *Neurospora crassa* concerned with NADH-oxidation. Relationship to a soluble NAD-reducing hydrogenase of *Alcaligenes eutrophus*. Biochim Biophys Acta 1090:133-138 [2]

Perez DM, Hwa J, Zhao MM, Porter J (1998) Molecular mechanisms of ligand binding and activation in alpha 1-adrenergic receptors. Adv Pharmacol. 42:398-403 [3]

Prodouz KN, Garrett RH (1981) *Neurospora crassa* NAD(P)H-nitrite reductase. Studies on its composition and structure. J Biol Chem 256:9711-9717 [2]

Rae JL, Shepard AR (2000) Kv3.3 potassium channels in lens epithelium and corneal endothelium. Exp Eye Res 70:339-348 [4]

Ramirez-Latorre J, Yu CR, Qu X, Perin F, Karlin A, Role L (1996) Functional contributions of alpha5 subunit to neuronal acetylcholine receptor channels. Nature 380:347-351 [3]

Rasmusson AG, Heiser V, Irrgang KD, Brennicke A, Grohmann L (1998) Molecular characterisation of the 76 kDa iron-sulphur protein subunit of potato mitochondrial complex I. Plant Cell Physiol 39:373-381 [2]

Rasmusson RL, Morales MJ, Wang S, Liu S, Campbell DL, Brahmajothi MV, Strauss HC (1998) Inactivation of voltage-gated cardiac K+ channels. Circ Res 82:739-750 [4]

Ravichandran KG, Boddupalli SS, Hasermann CA, Peterson JA, Deisenhofer J (1993) Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. Science 261:731-736 [2]

Reeves DC, Lummis SC (2002) The molecular basis of the structure and function of the 5-HT3 receptor: a model ligand-gated ion channel (review). Mol Membr Biol

19:11-26 [3]

Regulski M, Tully T (1995) Molecular and biochemical characterization of dNOS: a *Drosophila* Ca^{2+}/calmodulin-dependent nitric oxide synthase. Proc Natl Acad Sci U S A 92:9072-9076 [2]

Rettig J, Heinemann SH, Wunder F, Lorra C, Parcej DN, Dolly JO, Pongs O (1994) Inactivation properties of voltage-gated K+ channels altered by presence of beta-subunit. Nature 369:289-294 [4]

Rettig J, Wunder F, Stocker M, Lichtinghagen R, Mastiaux F, Beckh S, Kues W, Pedarzani P, Schroter KH, Ruppersberg JP, Veh R, Pongs O (1992) Characterization of a Shaw-related potassium channel family in rat brain. EMBO J 11:2473-2486 [4]

Reyes F, Roldan MD, Klipp W, Castillo F, Moreno-Vivian C (1996) Isolation of periplasmic nitrate reductase genes from *Rhodobacter sphaeroides* DSM 158: structural and functional differences among prokaryotic nitrate reductases. Mol Microbiol 19:1307-1318 [2]

Reynolds JA, Karlin A (1978) Molecular weight in detergent solution of acetylcholine receptor from *Torpedo californica*. Biochemistry 17:2035-2038 [3]

Roth BL, Willins DL, Kristiansen K, Kroeze WK (1998) 5-Hydroxytryptamine2-family receptors (5-hydroxytryptamine2A, 5-hydroxytryptamine2B, 5-hydroxytryptamine2C): where structure meets function. Pharmacol Ther 79:231-257 [3]

Rothery RA, Trieber CA, Weiner JH (1999) Interactions between the molybdenum cofactor and iron-sulfur clusters of *Escherichia coli* dimethylsulfoxide reductase. J Biol Chem 274:13002-13009 [2]

Rothery RA, Weiner JH (1996) Interaction of an engineered [3Fe-4S] cluster with a menaquinol binding site of *Escherichia coli* DMSO reductase. Biochemistry 35:3247-3257 [2]

Rubio LM, Herrero A, Flores E (1996) A cyanobacterial *narB* gene encodes a ferredoxin-dependent nitrate reductase. Plant Mol Biol 30:845-850 [2]

Sack JS, Saper MA, Quiocho FA (1989a) Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. J Mol Biol 206:171-191 [3]

Sack JS, Trakhanov SD, Tsigannik IH, Quiocho FA (1989b) Structure of the L-leucine-binding protein refined at 2.4 A resolution and comparison with the Leu/Ile/Val-binding protein structure. J Mol Biol 206:193-207 [3]

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425 [2, 3, 4]

Saraste M, Castresana J (1994) Cytochrome oxidase evolved by tinkering with denitrification enzymes. FEBS Lett 341:1-4 [1, 2]

Sartania N, Strange PG (1999) Role of conserved serine residues in the interaction of agonists with D3 dopamine receptors. J Neurochem 72:2621-2624 [3]

Sauter M, Bohm R, Bock A (1992) Mutational analysis of the operon (*hyc*) determining hydrogenase 3 formation in *Escherichia coli*. Mol Microbiol 6:1523-1532 [2]

Schapira M, Abagyan R, Totrov M (2002) Structural model of nicotinic acetylcholine receptor isotypes bound to acetylcholine and nicotine. BMC Struct Biol 2:1 [3]

Schindelin H, Kisker C, Hilton J, Rajagopalan KV, Rees DC (1996) Crystal structure of DMSO reductase: redox-linked changes in molybdopterin coordination. Science 272:1615-1621 [2]

Schneider F, Lowe J, Huber R, Schindelin H, Kisker C, Knablein J (1996) Crystal structure of dimethyl sulfoxide reductase from *Rhodobacter capsulatus* at 1.88 A resolution. J Mol Biol 263:53-69 [2]

Seeber M, De Benedetti PG, Fanelli F (2003) Molecular dynamics simulations of the ligand-induced chemical information transfer in the 5-HT_{1A} receptor. J Chem Inf Comput Sci 43:1520-1531 [3]

Serre L, Vellieux FM, Medina M, Gomez-Moreno C, Fontecilla-Camps JC, Frey M (1996) X-ray structure of the ferredoxin:$NADP^+$ reductase from the cyanobacterium *Anabaena* PCC 7119 at 1.8 A resolution, and crystallographic studies of $NADP^+$ binding at 2.25 A resolution. J Mol Biol 263:20-39 [2]

Sevrioukova I, Shaffer C, Ballou DP, Peterson JA (1996) Equilibrium and transient state spectrophotometric studies of the mechanism of reduction of the flavoprotein domain of P450BM-3. Biochemistry 35:7058-7068 [2]

Sevrioukova IF, Hazzard JT, Tollin G, Poulos TL (1999a) The FMN to heme electron transfer in cytochrome P450BM-3. Effect of chemical modification of cysteines engineered at the FMN-heme domain interaction site. J Biol Chem 274:36097-36106 [2]

Sevrioukova IF, Li H, Zhang H, Peterson JA, Poulos TL (1999b) Structure of a cytochrome P450-redox partner electron-transfer complex. Proc Natl Acad Sci USA 96:1863-1868 [2]

Shaw JP, Harayama S (1992) Purification and characterisation of the NADH:acceptor reductase component of xylene monooxygenase encoded by the TOL plasmid pWW0 of *Pseudomonas putida* mt-2. Eur J Biochem 209:51-61 [2]

Shelton PA, Dart C (1996) The use of the rat LAP LCR and promoter for the high-level

constitutive expression of K+ channel cDNAs in a rat liver cell line. FEBS Lett 391:282-286 [4]

Shibata R, Nakahira K, Shibasaki K, Wakazono Y, Imoto K, Ikenaka K (2000) A-type K+ current mediated by the Kv4 channel regulates the generation of action potential in developing cerebellar granule cells. J Neurosci 20:4145-55 [4]

Schmieden V, Kuhse J, Betz H (1993) Mutation of glycine receptor subunit creates beta-alanine receptor responsive to GABA. Science 262:256-258 [3]

Shuber AP, Orr EC, Recny MA, Schendel PF, May HD, Schauer NL, Ferry JG (1986) Cloning, expression, and nucleotide sequence of the formate dehydrogenase genes from *Methanobacterium formicicum*. J Biol Chem 261:12942-12947 [2]

Siddhanta U, Presta A, Fan B, Wolan D, Rousseau DL, Stuehr DJ (1998) Domain swapping in inducible nitric-oxide synthase. Electron transfer occurs between flavin and heme groups located on adjacent subunits in the dimer. J Biol Chem 273:18950-18958 [2]

Siddiqui RA, Warnecke-Eberz U, Hengsberger A, Schneider B, Kostka S, Friedrich B (1993) Structure and function of a periplasmic nitrate reductase in *Alcaligenes eutrophus* H16. J Bacteriol 175:5867-5876 [2]

Siegel LM, Murphy MJ, Kamin H (1973) Reduced nicotinamide adenine dinucleotide phosphate-sulfite reductase of enterobacteria. I. The *Escherichia coli* hemoflavoprotein: molecular parameters and prosthetic groups. J Biol Chem 248:251-264 [2]

Simon MJ, Osslund TD, Saunders R, Ensley BD, Suggs S, Harcourt A, Suen WC, Cruden DL, Gibson DT, Zylstra GJ (1993) Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4.

Gene 127:31-37 [2]

Skipper L, Campbell WH, Mertens JA, Lowe DJ (2001) Pre-steady-state kinetic analysis of recombinant *Arabidopsis* NADH:nitrate reductase: rate-limiting processes in catalysis. J Biol Chem 276:26995-27002 [2]

Sridhar Prasad G, Kresge N, Muhlberg AB, Shaw A, Jung YS, Burgess BK, Stout CD (1998) The crystal structure of NADPH:ferredoxin reductase from *Azotobacter vinelandii*. Protein Sci 7:2541-2549 [2]

Stewart LJ, Bailey S, Bennett B, Charnock JM, Garner CD, McAlpine AS (2000) Dimethylsulfoxide reductase: an enzyme capable of catalysis with either molybdenum or tungsten at the active site. J Mol Biol 299:593-600 [2]

Strader CD, Candelore MR, Hill WS, Sigal IS, Dixon RA (1989) Identification of two serine residues involved in agonist activation of the beta-adrenergic receptor. J Biol Chem 264:13572-13578 [3]

Strader CD, Sigal IS, Candelore MR, Rands E, Hill WS, Dixon RA (1988) Conserved aspartic acid residues 79 and 113 of the beta-adrenergic receptor have different roles in receptor function. J Biol Chem 263:10267-10271 [3]

Steward LJ, Boess FG, Steele JA, Liu D, Wong N, Martin IL (2000) Importance of phenylalanine 107 in agonist recognition by the 5-hydroxytryptamine$_{3A}$ receptor. Mol Pharmacol 57:1249-1255 [3]

Strittmatter P, Kittler JM, Coghill JE, Ozols J (1992) Characterization of lysyl residues of NADH-cytochrome b_5 reductase implicated in charge-pairing with active-site carboxyl residues of cytochrome b_5 by site-directed mutagenesis of an expression vector for the flavoprotein. J Biol Chem 267:2519-23 [2]

Stuhmer W, Ruppersberg JP, Schroter KH, Sakmann B, Stocker M, Giese KP, Perschke

A, Baumann A, Pongs O (1989) Molecular basis of functional diversity of voltage-gated potassium channels in mammalian brain. EMBO J 8:3235-3244 [4]

Sun Y, Olson R, Horning M, Armstrong N, Mayer M, Gouaux E (2002) Mechanism of glutamate receptor desensitization. Nature 417:245-253 [3]

Sun YJ, Rose J, Wang BC, Hsiao CD (1998) The structure of glutamine-binding protein complexed with glutamine at 1.94 A resolution: comparisons with other amino acid binding proteins. J Mol Biol 278:219-229 [3]

Tan J, Cowan JA (1991) Enzymatic redox chemistry: a proposed reaction pathway for the six-electron reduction of SO_3^{2-} to S^{2-} by the assimilatory-type sulfite reductase from *Desulfovibrio vulgaris* (Hildenborough). Biochemistry 30:8910-8917 [2]

Tanabe T, Takeshima H, Mikami A, Flockerzi V, Takahashi H, Kangawa K, Kojima M, Matsuo H, Hirose T, Numa S (1987) Primary structure of the receptor for calcium channel blockers from skeletal muscle. Nature 328:313-318 [4]

Tatusova TA, Madden TL (1999) BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol Lett 174:247-250 [3]

Tempel BL, Papazian DM, Schwarz TL, Jan YN, Jan LY (1987) Sequence of a probable potassium channel component encoded at Shaker locus of Drosophila. Science. 237:770-775 [4]

Temple CA, George GN, Hilton JC, George MJ, Prince RC, Barber MJ, Rajagopalan KV (2000) Structure of the molybdenum site of *Rhodobacter sphaeroides* biotin sulfoxide reductase. Biochemistry 39:4046-4052 [2]

ter Laak AM, Timmerman H, Leurs R, Nederkoorn PH, Smit MJ, Donne-Op den Kelder GM (1995) Modelling and mutation studies on the histamine H1-receptor agonist binding site reveal different binding modes for H1-agonists: Asp116 (TM3) has a

constitutive role in receptor stimulation. J Comput Aided Mol Des 9:319-330 [3]

Thomas G, Potter L, Cole JA (1999) The periplasmic nitrate reductase from *Escherichia coli*: a heterodimeric molybdoprotein with a double-arginine signal sequence and an unusual leader peptide cleavage site. FEMS Microbiol Lett 174:167-171 [2]

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876-4882 [2, 3, 4]

Tsaur ML, Chou CC, Shih YH, Wang HL (1997) Cloning, expression and CNS distribution of Kv4.3, an A-type K+ channel alpha subunit. FEBS Lett 400:215-220 [4]

Tseng-Crank JC, Tseng GN, Schwartz A, Tanouye MA (1990) Molecular cloning and functional expression of a potassium channel cDNA isolated from a rat cardiac library. FEBS Lett 268:63-68 [4]

Trieber CA, Rothery RA, Weiner JH (1996) Engineering a novel iron-sulfur cluster into the catalytic subunit of *Escherichia coli* dimethyl-sulfoxide reductase. J Biol Chem 271:4620-4626 [2]

Unkles SE, Campbell EI, Punt PJ, Hawker KL, Contreras R, Hawkins AR, Van den Hondel CA, Kinghorn JR (1992) The *Aspergillus niger niaD* gene encoding nitrate reductase: upstream nucleotide and amino acid sequence comparisons. Gene 111:149-155 [2]

Uveges AJ, Kowal D, Zhang Y, Spangler TB, Dunlop J, Semus S, Jones PG (2002) The role of transmembrane helix 5 in agonist binding to the human H3 receptor. J Pharmacol Exp Ther 301:451-458 [3]

Vafa B, Lewis TM, Cunningham AM, Jacques P, Lynch JW, Schofield PR (1999)

Identification of a new ligand binding domain in the alpha1 subunit of the inhibitory glycine receptor. J Neurochem 73:2158-2166 [3]

Vaidehi N, Floriano WB, Trabanino R, Hall SE, Freddolino P, Choi EJ, Zamanakos G, Goddard WA 3rd (2002) Prediction of structure and function of G protein-coupled receptors. Proc Natl Acad Sci U S A 99:12622-12627 [3]

van den Akker F, Zhang X, Miyagi M, Huo X, Misono KS, Yee VC (2000) Structure of the dimerized hormone-binding domain of a guanylyl-cyclase-coupled receptor. Nature 406:101-104 [3]

Vandenberg RJ, French CR, Barry PH, Shine J, Schofield PR (1992a) Antagonism of ligand-gated ion channel receptors: two domains of the glycine receptor alpha subunit form the strychnine-binding site. Proc Natl Acad Sci U S A 89:1765-1769 [3]

Vandenberg RJ, Handford CA, Schofield PR (1992b) Distinct agonist- and antagonist-binding sites on the glycine receptor. Neuron 9:491-496 [3]

van den Berg WA, Hagen WR, van Dongen WM (2000) The hybrid-cluster protein ('prismane protein') from *Escherichia coli*. Characterization of the hybrid-cluster protein, redox properties of the [2Fe-2S] and [4Fe-2S-2O] clusters and identification of an associated NADH oxidoreductase containing FAD and [2Fe-2S]. Eur J Biochem 267:666-676 [2]

van Thor JJ, Geerlings TH, Matthijs HC, Hellingwerf KJ (1999) Kinetic evidence for the PsaE-dependent transient ternary complex photosystem I/Ferredoxin/Ferredoxin:$NADP^+$ reductase in a cyanobacterium. Biochemistry 38:12735-12746 [2]

Varshney A, S K, Mathew MK (2002) Modulation of voltage sensitivity by N-terminal

cytoplasmic residues in human Kv1.2 channels. Eur Biophys J 31:365-372 [4]

Vega JM (1976) A reduced pyridine nucleotides-diaphorase activity associated to the assimilatory nitrite reductase complex from *Neurospora crassa*. Arch Microbiol 109:237-242 [2]

Vega-Saenz de Miera E, Rudy B (1992) Modulation of K+ channels by hydrogen peroxide. Biochem Biophys Res Commun 186:1681-1687 [4]

Venturi V, Zennaro F, Degrassi G, Okeke BC, Bruschi CV (1998) Genetics of ferulic acid bioconversion to protocatechuic acid in plant-growth-promoting *Pseudomonas putida* WCS358. Microbiology 144:965-973 [2]

Vernallis AB, Conroy WG, Berg DK (1993) Neurons assemble acetylcholine receptors with as many as three kinds of subunits while maintaining subunit segregation among receptor subtypes. Neuron 10:451-464 [3]

Wafford KA, Kathoria M, Bain CJ, Marshall G, Le Bourdelles B, Kemp JA, Whiting PJ (1995) Identification of amino acids in the N-methyl-D-aspartate receptor NR1 subunit that contribute to the glycine binding site. Mol Pharmacol 47:374-380 [3]

Wang CD, Buck MA, Fraser CM (1991) Site-directed mutagenesis of alpha 2A-adrenergic receptors: identification of amino acids involved in ligand binding and receptor activation by agonists. Mol Pharmacol 40:168-179 [3]

Ward SD, Curtis CA, Hulme EC (1999) Alanine-scanning mutagenesis of transmembrane domain 6 of the M_1 muscarinic acetylcholine receptor suggests that Tyr381 plays key roles in receptor function. Mol Pharmacol 56:1031-1041 [3]

Watanabe M, Maemura K, Kanbara K, Tamayama T, Hayasaki H (2002) GABA and GABA receptors in the central nervous system and other organs. Int Rev Cytol 213:1-47 [3]

Wedel B, Humbert P, Harteneck C, Foerster J, Malkewitz J, Bohme E, Schultz G, Koesling D (1994) Mutation of His-105 in the beta 1 subunit yields a nitric oxide-insensitive form of soluble guanylyl cyclase. Proc Natl Acad Sci U S A 91:2592-2596 [3]

Wei A, Covarrubias M, Butler A, Baker K, Pak M, Salkoff L (1990) K+ current diversity is produced by an extended gene family conserved in Drosophila and mouse. Science 248:599-603 [4]

Weiser M, Vega-Saenz de Miera E, Kentros C, Moreno H, Franzen L, Hillman D, Baker H, Rudy B (1994) Differential expression of Shaw-related K+ channels in the rat central nervous system. J Neurosci 14:949-972 [4]

Werkman TR, Kawamura T, Yokoyama S, Higashida H, Rogawski MA (1992) Charybdotoxin, dendrotoxin and mast cell degranulating peptide block the voltage-activated K+ current of fibroblast cells stably transfected with NGK1 (Kv1.2) K+ channel complementary DNA. Neuroscience 50:935-946 [4]

White JH, Wise A, Main MJ, Green A, Fraser NJ, Disney GH, Barnes AA, Emson P, Foord SM, Marshall FH (1998) Heterodimerization is required for the formation of a functional $GABA_B$ receptor. Nature 396:679-682 [3]

Wieczorek H, Brown D, Grinstein S, Ehrenfeld J, Harvey WR (1999) Animal plasma membrane energization by proton-motive V-ATPases. Bioessays 21:637-648 [2]

Wieland K, Laak AM, Smit MJ, Kuhne R, Timmerman H, Leurs R (1999) Mutational analysis of the antagonist-binding site of the histamine H_1 receptor. J Biol Chem 274:29994-30000 [3]

Wymore RS, Korenberg JR, Kinoshita KD, Aiyar J, Coyne C, Chen XN, Hustad CM, Copeland NG, Gutman GA, Jenkins NA, Chandy KG (1994) Genomic organization,

nucleotide sequence, biophysical properties, and localization of the voltage-gated K+ channel gene KCNA4/Kv1.4 to mouse chromosome 2/human 11p14 and mapping of KCNC1/Kv3.1 to mouse 7/human 11p14.3-p15.2 and KCNA1/Kv1.1 to human 12p13. Genomics 20:191-202 [4]

Xu X, Matsuno-Yagi A, Yagi T (1992) Structural features of the 66-kDa subunit of the energy-transducing NADH-ubiquinone oxidoreductase (NDH-1) of *Paracoccus denitrificans*. Arch Biochem Biophys 296:40-48 [2]

Yabusaki Y, Murakami H, Ohkawa H (1988) Primary structure of *Saccharomyces cerevisiae* NADPH-cytochrome P450 reductase deduced from nucleotide sequence of its cloned gene. J Biochem 103:1004-1010 [2]

Yen KM, Karl MR (1992) Identification of a new gene, *tmoF*, in the *Pseudomonas mendocina* KR1 gene cluster encoding toluene-4-monooxygenase. J Bacteriol 174:7253-7261 [2]

Yeom H, Sligar SG, Li H, Poulos TL, Fulco AJ (1995) The role of Thr268 in oxygen activation of cytochrome P450BM-3. Biochemistry 34:14733-14740 [2]

Zenno S, Hattori M, Misumi Y, Yubisui T, Sakaki Y (1990) Molecular cloning of a cDNA encoding rat NADH-cytochrome b_5 reductase and the corresponding gene. J Biochem 107:810-816 [2]

Zerr P, Adelman JP, Maylie J (1998) Characterization of three episodic ataxia mutations in the human Kv1.1 potassium channel. FEBS Lett 431:461-464 [4]

Zhang Z, Qiu W, Quinn SJ, Conigrave AD, Brown EM, Bai M (2002) Three adjacent serines in the extracellular domains of the CaR are required for L-amino acid-mediated potentiation of receptor function. J Biol Chem 277:33727-33735 [3]

Zhu H, Riggs AF (1992) Yeast flavohemoglobin is an ancient protein related to globins

and a reductase family. Proc Natl Acad Sci U S A 89:5015-5019 [2]

Zhu XR, Netzer R, Bohlke K, Liu Q, Pongs O (1999) Structural and functional characterization of Kv6.2 a new gamma-subunit of voltage-gated potassium channel. Receptors Channels 6:337-350 [4]

Zumft WG, Dreusch A, Lochelt S, Cuypers H, Friedrich B, Schneider B (1992) Derived amino acid sequences of the *nosZ* gene (respiratory N_2O reductase) from *Alcaligenes eutrophus*, *Pseudomonas aeruginosa* and *Pseudomonas stutzeri* reveal potential copper-binding residues. Implications for the CuA site of N_2O reductase and cytochrome-c oxidase. Eur J Biochem 208:31-40 [1, 2]

Acknowledgments

I am grateful to Dr. Hiromi Suzuki of Meiji University for helping me in the initial part of the study in chapter 2. I appreciate Dr. Takeshi Kawabata of NAIST for advising me the method of PSI-BLAST homology search method and helping me to conduct it in chapter 2. Members in Saitou laboratory kindly gave me good advices about all parts of this study, and I'm thankful to them. Most of all, I appreciate Dr. Saitou to teach me and give me good advices in all parts.