We anticipate that further experimental investigation of the specific role of each of the five classes of Brix domain proteins in eukaryote model organisms, especially in yeast, will yield valuable new insights about ribosome synthesis and be key to our understanding of the ribosome biogenesis pathway.

Acknowledgements
We are grateful to E. Bogengruber, M. Breitenbach, F.M. Jantsch and G. Lepperdinger for supplying experimental data on sequence and function of Brix (AF319877) and yol077c before publication and for extensive discussion of the Brix sequence analysis results. This research was supported by Boehringer-Ingelheim International.

References

1 Kressler, D. *et al.* (1999) Protein *trans*-acting factors involved in ribosome biogenesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 19, 7897–7912
2 Altschul, S.F. *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25, 3389–3402
3 PFAM: http://pfam.wustl.edu
4 Migeon, J.C. *et al.* (1999) Cloning and characterization of Peter Pan, a novel *Drosophila* gene required for larval growth. *Mol. Biol. Cell* 10, 1733–1744
5 Mayer, C. *et al.* (2001) The archaeal homologue of the IMP4 protein, a eukaryotic U3 snoRNP component. *Trends Biochem. Sci.* 26, 143–144
6 Wootton, J.C. and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* 266, 554–571
7 Suarez-Huerta, N. *et al.* (2000) Cloning, genomic organization, and tissue distribution of human Ssf-1. *Biochem. Biophys. Res. Commun.* 275, 37–42
8 Yu, Y. and Hirsch, J.P. (1995) An essential gene pair in *Saccharomyces cerevisiae* with a potential role in mating. *DNA Cell Biol.* 14, 411–418
9 Kim, J. and Hirsch, J.P. (1998) A nucleolar protein that affects mating efficiency in *Saccharomyces cerevisiae* by altering the morphological response to pheromone. *Genetics* 149, 795–805
10 Nathan, D.F. *et al.* (1999) Identification of Ssf1, Cns1, and Hch1 as multicopy suppressors of a *Saccharomyces cerevisiae* Hsp90 loss-of-function mutation. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1409–1414
11 Eurofan: http://www.mips.biochem.mpg.de/proj/yeast/search/code_search.htm
12 Gönczy, P. *et al.* (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* 408, 331–336
13 http://mpi-web.embl-heidelberg.de/dbScreen/
14 Lee, S.J. and Baserga, S.J. (1999) Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. *Mol. Cell. Biol.* 19, 5441–5452
15 Ding, D.Q. *et al.* (2000) Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic library. *Genes Cells* 5, 169–190
16 Cuff, J.A.and Barton, G.J. (2000) Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins* 40, 502–511

Frank Eisenhaber
Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, A-1030 Vienna, Rep. Austria.
e-mail: Frank.Eisenhaber@nt.imp.univie.ac.at

Christian Wechselberger
NIH-NCI, Tumor Growth Factor Section, Bldg. 10, R. 5B43, 10 Center Dr., MSC-1750, Bethesda, MD 20892, USA.

Günther Kreil
Institute of Molecular Biology, Austrian Academy of Sciences, Billrothstraße 11, A-5020 Salzburg, Austria.

A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems

Kay Hofmann and Laurent Falquet

Ubiquitination generally serves as a signal for targeting cytoplasmic and nuclear proteins to the proteasome for subsequent degradation. Recently, evidence has accumulated indicating that ubiquitination also plays an important role in targeting integral membrane proteins for degradation by the lytic vacuole or the lysosome. This article describes a conserved protein motif, based on a sequence of the proteasomal component Rpn10/S5a, that is known to recognize ubiquitin. The presence of this motif in Eps15, Epsin and HRS, proteins involved in ligand-activated receptor endocytosis and degradation, suggest a more general role in ubiquitin recognition.

Ubiquitin is a small protein, highly conserved among eukaryotes, that becomes covalently attached to both itself and a variety of cellular proteins[1,2]. The role of this ubiquitination is mostly to target proteins to the 26S proteasome degradation pathway[3]. In some cases, monoubiquitination (e.g. of histones) does not lead to degradation, but instead regulates other cellular processes such as chromatin remodeling[4]. Recently, several reports have described a role for monoubiquitination in a different pathway of protein degradation – the endocytosis and subsequent proteolysis of receptors and other transmembrane proteins by the vacuole or the lysosome[5,6]. According to the current model, the decisions about which protein is to be degraded at a specific time is made by the ubiquitination machinery, often in response to a prior event such as phosphorylation. Consequently, both the proteasome and the endocytosis machinery need a mechanism by which to faithfully recognize ubiquitinated proteins.

The 26S proteasome comprises two main particles: the 20S core proteasome and the 19S regulatory complex. Subunit S5a (also known as Rpn10) of the 19S regulator binds polyubiquitin chains and has a preference for chains containing four or more ubiquitin monomers. The ubiquitin-interacting region has been mapped to two short, related motifs that are found in all members of the S5a family[7]. Using these regions, which comprise ~20 residues, as a starting point, we searched for other potential ubiquitin-binding sequences. Specifically, we used a combination of iterative database searches with generalized profiles, and Hidden Markov Models (profile-HMMs)[8]. Only sequences that matched a profile or an HMM derived from previously established family members, with error probabilities of $p < 0.01$, were used for subsequent iteration cycles. After eight cycles, the sequence motif

  PII: S0968-0004(01)01835-7

converged to a set of proteins shown in Fig. 1, most of which had multiple copies of the motif. The recognizable part of the motif contains 20 residues, comprising a highly conserved Φ-x-x-Ala-x-x-x-Ser-x-x-Ac core, in which Φ denotes a large hydrophobic residue and 'Ac' denotes an acidic residue. This core region is preceded by a block of four preferentially acidic residues. Such a short sequence motif is unlikely to form an independent folding domain. Instead, based on the spacing of the conserved residues, the motif probably forms a short α helix that can be embedded into different protein folds.

As shown in Fig. 2, the ubiquitin-interacting motif (UIM) occurs in a wide variety of proteins, most of them having a modular architecture; that is, bearing multiple homology domains. In addition to the proteasomal S5a subunit (in which the region corresponding to the UIM is known to bind ubiquitin), the UIM is found in proteins either involved in ubiquitination and ubiquitin metabolism, or known to interact with ubiquitin-like modifiers. Among the UIM proteins are two different subgroups of the UBP family of deubiquitinating enzymes, one F-box protein, one family of HECT-containing ubiquitin-ligases (E3s) from plants, and several proteins containing ubiquitin-associated UBA and/or UBX domains[9]. By contrast, no ubiquitin-activating enzyme (E1) or ubiquitin-conjugating enzyme (E2) have been found to contain a UIM domain.

An observation of particular interest is the occurrence of UIMs in four classes of proteins involved in receptor endocytosis – the Eps15 subfamily of EH-domain proteins, the epsin subfamily of ENTH-domain proteins and two families of VHS-domain proteins, including the FYVE-finger proteins HRS and Vps27, and the SH3-domain proteins STAM and HBP. Eps15 is phosphorylated on Tyr850 by the ligand-activated epidermal growth factor (EGF) receptor and this phosphorylation is required for subsequent receptor endocytosis[10]. Furthermore, Eps15 binds to epsin, and both of these proteins interact with components of the endocytosis machinery, including clathrin and the AP-2 complex[11,12]. The fact that the phosphorylation site of Eps15 is immediately adjacent to a tandem UIM suggests that this motif might be involved in the regulated endocytosis of the EGF receptor. Liquid facets (lqf), an epsin from

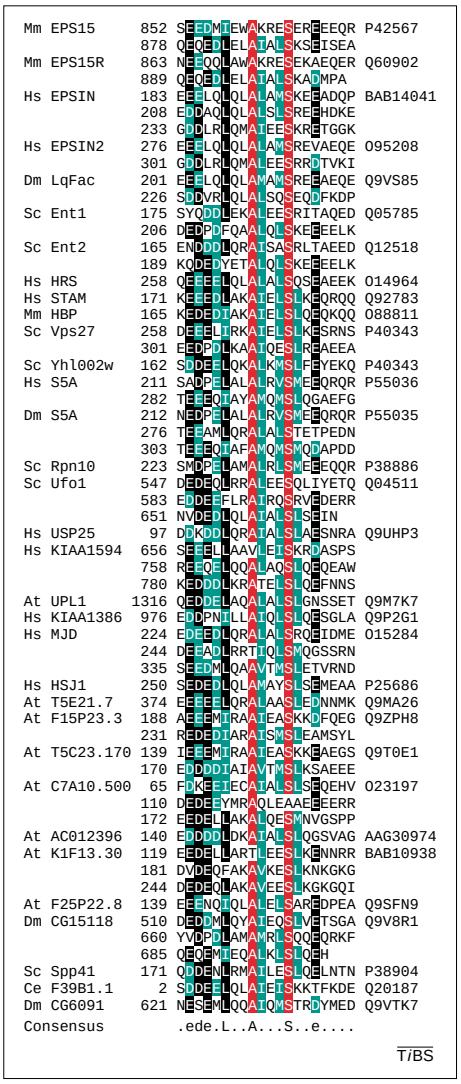

Fig. 1. Alignment of representative UIM motifs. Residues invariant or conservatively substituted in ≥50% of the sequences are printed on black and green backgrounds, respectively. The nearly invariant Ala and Ser positions are highlighted in red. For each sequence, the species is given in the leftmost column, followed by the protein name. Database accession numbers are given in the rightmost column and refer to the SwissProt/TrEMBL database. The bottom line indicates the consensus sequence with uppercase letters representing positions conserved in ≥50% of the sequences and lowercase letters representing visible conservations in <50% of the sequences. Abbreviations: Hs, *Homo sapiens*; Mm, *Mus musculus*; Dm, *Drosophila melanogaster*; At, *Arabidopsis thaliana*; Ce, *Caenorhabditis elegans*; Sc, *Saccharomyces cerevisiae*. This multiple sequence alignment (alignment number ALIGN_000115) has been deposited with the European Bioinformatics Institute (ftp://ftp.ebi.ac.uk/pub/databases/embl/aling/ALIGN_000115.dat).

Drosophila melanogaster, was identified in a genetic screen as a dominant enhancer of the fat facets (faf) mutant eye phenotype[13]. This relationship links the gene encoding lqf with the ubiquitin system because faf is a deubiquitinating enzyme. An additional link between Eps15 and ubiquitin recognition is provided by the yeast protein Ede1p, the closest homolog of mammalian Eps15 (Ref. 14): in the yeast protein, the UIM is replaced by a UBA domain, a homology domain known to bind ubiquitin[15].

Proteins with an N-terminal VHS domain can be divided into three subtypes on the basis of their domain organization: (1) Vps27–HRS-like (type A); (2) STAM–HBP-like (type B); and (3) other proteins (type C) (see Fig. 2, type C not shown). Several reports document an interaction between members of subfamilies A and B (Refs 16,17). These two types of VHS proteins become Tyr phosphorylated upon activation of different growth factor receptors. In budding yeast, Vps27p is involved in targeting proteins to the vacuole. Recently, a role of HRS in ligand-induced degradation of the EGF receptor (but not its endocytosis) has been demonstrated[18]. Two further observations have confirmed the connection between the ubiquitin and the endocytotic pathways with respect to VHS proteins: (1) HBP contains a UIM and binds to the mouse deubiquitinating enzyme UBPY (Ref. 19); and (2), in yeast, Vps27p interacts genetically, and colocalizes, with Doa4p, another deubiquitinating enzyme[20].

Finally, the functions of several UIM-containing proteins suggest a role of this motif in the pathogenesis of neurodegeneration. Machado–Joseph Disease (MJD) is an autosomal dominant neurodegenerative disease caused by expansion of the polyGln tract near the C-terminus of MJD1 (also known as ataxin-3). A hallmark of polyGln-linked ataxias are intranuclear inclusions, which contain ubiquitin and, at least in the case of MJD, components of the proteasome[21]. The ataxin-3 protein contains a tandem UIM upstream of the polyGln sequence and, in one splice variant, an additional UIM downstream of the expansion. Ataxin-3 has been reported to bind to HHR23, a human homolog of Rad23p that contains ubiquitin-related and -associated domains[22]. However, the UIM region of ataxin-3 is not required for this interaction. In addition, genetic screens in *D. melanogaster* have recently provided evidence that the pathology associated with polyGln expansion can be suppressed by increased expression of J-domain-containing chaperonins[23,24]; interestingly, at least one human DnaJ homolog (HSJ1) contains a recognizable UIM motif. A further indirect link between ubiquitination and neurodegeneration is provided by KIAA1386, an uncharacterized

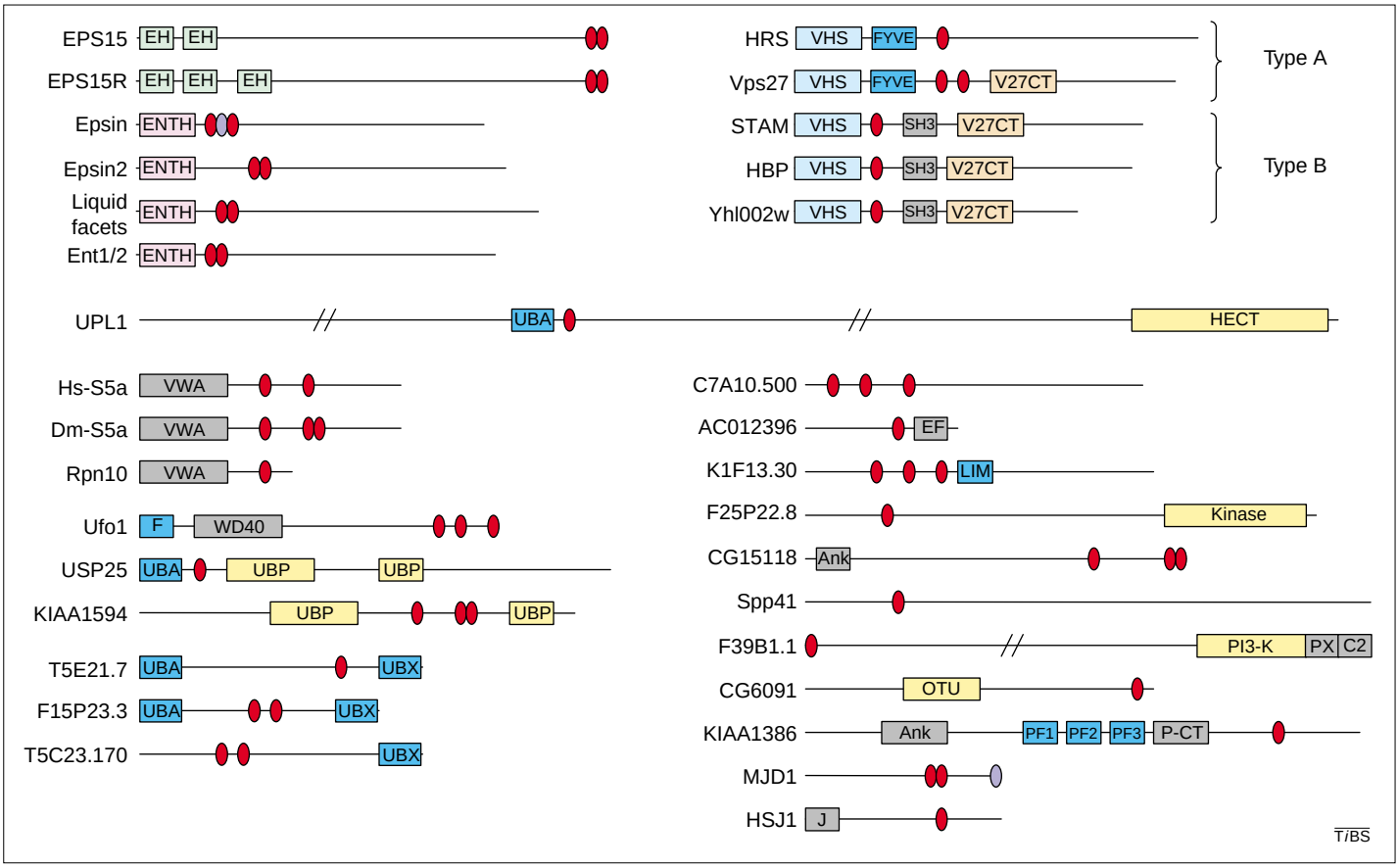

Fig. 2. Schematic diagram illustrating the domains of representative proteins harboring the UIM motif. All proteins (except for UPL1 and F39B1.1) are drawn approximately to scale. UIM motifs are indicated by red ovals, with a subset of UIMs occurring only in selected splice variants being colored purple. Other domains are drawn as colored rectangles. Yellow indicates catalytically active domains, cyan indicates domains involved in ubiquitination and green indicates various types of complex zinc-finger domains resembling the RING-finger domain. Protein names are the same as in Fig. 1, in which the corresponding species and accession numbers are also given. Abbreviations: Ank, ankyrin repeat region; C2, C2-domain; EF, EF-hand Ca^{2+}-binding motif; EH, Eps15-homology domain; ENTH, epsin N-terminal homology domain; F, F-box domain; FYVE, FYVE-finger domain; HECT, ubiquitin transferase catalytic domain; J, dnaJ chaperone domain; kinase, protein kinase catalytic domain; LIM, LIM-type Zn-finger domain; OTU, OTU-protease catalytic domain; P-CT, parkin C-terminal domain; PF1–3, parkin-type complex Zn-finger domains 1–3; PI3-K, PI3-kinase catalytic domain; PX, phox-domain; SH3, Src-homology domain 3; UBA, ubiquitin-associated domain; UBP, bipartite ubiquitin protease catalytic domain; UBX; ubiquitin-related X domain; V27CT, Vps27 C-terminal homology domain; VHS, Vps27–HRS–Stam domain; VWA, von Willebrand factor type A domain; WD40, WD40-repeat region.

human protein harboring a UIM next to a region with extensive similarity to parkin, a putative ubiquitin ligase whose deficiency is associated with hereditary Parkinson disease.

Because the UIM motif is very short, it cannot be guaranteed that all biologically meaningful instances are identified by our profile methods. Several potentially interesting proteins reach scores slightly below the significance threshold, including the DNA-repair protein XPG, the RING-finger protein MAT1 and members of the cyclin D3 family. We hypothesize that, similar to the UBA domain, the UIM might be a general ubiquitin-binding motif. Although this role is well documented for the UIM in the proteasomal S5a subunit, a similar function for the UIM present in the factors involved in endocytosis and lysosomal degradation remains to be established.

References

1 Hochstrasser, M. (1996) Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* 30, 405–439
2 Ciechanover, A. *et al.* (2000) Ubiquitin-mediated proteolysis: biological regulation via destruction. *Bioessays* 22, 442–451
3 Voges, D. *et al.* (1999) The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* 68, 1015–1068
4 Nickel, B.E. *et al.* (1989) Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochemistry* 28, 958–963
5 Hicke, L. (1999) Gettin' down with ubiquitin: turning off cell-surface receptors, transporters and channels. *Trends Cell Biol.* 9, 107–112
6 Shih, S.C. *et al.* (2000) Monoubiquitin carries a novel internalization signal that is appended to activated receptors. *EMBO J.* 19, 187–198
7 Young, P. *et al.* (1998) Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a. *J. Biol. Chem.* 273, 5461–5467
8 Hofmann, K. (2000) Sensitive protein comparisons with profiles and HMMs. *Briefings in Bioinformatics* 1, 167–178
9 Hofmann, K. and Bucher, P. (1996) The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends Biochem. Sci.* 21, 172–173
10 Confalonieri, S. *et al.* (2000) Tyrosine phosphorylation of Eps15 is required for ligand-regulated, but not constitutive, endocytosis. *J. Cell Biol.* 150, 905–912
11 Torrisi, M.R. *et al.* (1999) Eps15 is recruited to the plasma membrane upon epidermal growth factor receptor activation and localizes to components of the endocytic pathway during receptor internalization; *Mol. Biol. Cell* 10, 417–434
12 Drake, M.T. *et al.* (2000) Epsin binds to clathrin by associating directly with the clathrin-terminal domain. Evidence for cooperative binding through two discrete sites. *J. Biol. Chem.* 275, 6479–6489
13 Cadavid, A.L. *et al.* (2000) The function of the *Drosophila* fat facets deubiquitinating enzyme in limiting photoreceptor cell number is intimately associated with endocytosis. *Development* 127, 1727–1736
14 Gagny, B. *et al.* (2000) A novel EH domain protein *of Saccharomyces cerevisiae*, Ede1p, involved in endocytosis. *J. Cell Sci.* 113, 3309–3319
15 Vadlamudi, R.K. *et al.* (1996) p62, a phosphotyrosine-independent ligand of the SH2 domain of p56lck, belongs to a new class of ubiquitin-binding proteins. *J. Biol. Chem.* 271, 20235–20237

16 Komada, M. and Kitamura, N. (1995) Growth factor-induced tyrosine phosphorylation of Hrs, a novel 115-kilodalton protein with a structurally conserved putative zinc finger domain. *Mol. Cell. Biol.* 15, 6213–6221
17 Asao, H. *et al.* (1997) Hrs is associated with STAM, a signal-transducing adaptor molecule. Its suppressive effect on cytokine-induced cell growth. *J. Biol. Chem.* 272, 32785–32791
18 Chin, L. *et al.* Hrs interacts with sorting Nexin 1 and regulates degradation of epidermal growth factor receptor. *J. Biol. Chem.* (in press)
19 Kato, M. *et al.* (2000) A deubiquitinating enzyme UBPY interacts with the Src homology 3 domain of Hrs-binding protein via a novel binding motif PX(V/I)(D/N)RXXKP. *J. Biol. Chem.* 275, 37481–37487
20 Amerik, A.Y. *et al.* (2000) The Doa4 deubiquitinating enzyme is functionally linked to the vacuolar protein-sorting and endocytic pathways. *Mol. Biol. Cell* 11, 3365–3380
21 Chai, Y. *et al.* (1999) Evidence for proteasome involvement in polyglutamine disease: localization to nuclear inclusions in SCA3/MJD and suppression of polyglutamine aggregation *in vitro*. *Hum. Mol. Genet.* 8, 673–682
22 Wang, G.H. *et al.* (2000) Ataxin-3, the *MJD1* gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, HHR23A and HHR23B. *Hum. Mol. Genet.* 9, 1795–1803
23 Fernandez-Funez, P. *et al.* (2000) Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* 408, 101–106
24 Kazemi-Esfarjani, P. and Benzer, S. (2000) Genetic suppression of polyglutamine toxicity in *Drosophila*. *Science* 287, 1837–1840

Kay Hofmann
Bioinformatics and Gene Discovery Group, MEMOREC Stoffel GmbH, Stöckheimer Weg 1, D-50829 Köln, Germany.
e-mail: kay.hofmann@memorec.com

Laurent Falquet
Swiss Institute of Bioinformatics, Chemin des Boveresses 155, CH-1066 Epalinges, Switzerland.

Meeting Report

Membrane proteins on the move

Emma K. Wilson

The Keystone Symposium on Membrane protein structure/function relationships was held on 5–11 March 2001 in Tahoe City, California, USA.

Keystone Symposia have now been going strong for 30 years! The first conference was held at Squaw Valley (California, USA) in March 1972 and concentrated on membrane structure and function. The final speaker at that symposium was Ronald Kaback, who described his pioneering work on sugar and amino-acid transport. Now, Kaback has teamed-up with Ernest Wright to organize this superb conference on the structure–function relationships of membrane proteins. The 150 or so delegates at this meeting were treated to a broad selection of talks covering six main themes: general approaches, membrane protein structure, ion channels, transporters, ligand-gated channels and non-channel receptors.

Structural determination of membrane proteins

Although integral membrane proteins constitute approximately one-third of all gene products, their structural elucidation remains extremely challenging. Even in the era of structural genomics, only a handful of membrane protein structures have been solved to below 3 Å. This low 'success' rate reflects inherent difficulties with expression, purification and crystallization of membrane proteins. Presently, the only way to obtain <3 Å resolution data is by X-ray or electron diffraction and, generally, the rate-limiting step is finding crystallization conditions for the detergent-solubilized protein. Although there are several tricks to encourage crystal growth, for example the use of monoclonal antibody fragments, it is still very much a 'black art'. However, recently described 3D-crystallization techniques show considerable promise. Ehud Landau (University of Texas, Galveston, TX, USA) described lipid cubic phase crystallization (recently used to solve structures of bacteriorhodopsin), a novel method using membrane-like bicontinuous lipids that extend in three dimensions and embed the proteins. The advantages of this method are that the proteins are in their natural environment, can diffuse freely in the bilayer and are often more stable than in detergent micelles.

When good quality 3D crystals cannot be obtained, 2D crystals can sometimes be grown more easily, allowing electron diffraction studies by electron crystallography. This technique was pioneered by studies on bacteriorhodopsin and has recently revealed the architecture of aquaporins as discussed by Andreas Engel (University of Basel, Switzerland). Electron crystallography is becoming more widespread and has been used to visualize channel proteins, often confirming quaternary structures that have been postulated from biochemical studies. However, 'seeing' a membrane protein for the first time can produce surprises; for example, Chris Miller (Brandeis University, Waltham, MA, USA) explained that the dimeric bacterial CLC-type chloride channel clearly contains four visible holes, prompting the question: do each pair of holes represent different regions of the same functional pore?

Cysteine chemistry

The title of the conference includes the words 'structure' and 'function' and the speakers reflected this emphasis. Numerous talks highlighted the functional information obtained by biochemical and biophysical methods. A technique that addresses both protein topology and dynamics is the substituted cysteine accessibility method (SCAM), which takes advantage of cysteine thiol chemistry. Specific amino acids are mutated to Cys and their accessibility tested with thiol-reacting chemicals. Residues accessible from the extracellular side can be modified by both membrane-impermeable and -permeable reagents, whereas residues on the cytoplasmic side are only modified by the latter; residues that are completely buried will not react at all. Ernest Wright (University of California-Los Angeles, CA, USA) described how SCAM has been used extensively to study the sodium–glucose co-transporter. The ligand-induced conformational change of this protein has been probed and Cys residues located in transmembrane helices 10–13 become accessable during channel gating, implying large-scale helical movements of ~10 Å. Arthur Karlin (Columbia University, New

 0968-0004/01/$ – see front matter PII: S0968-0004(01)01873-4