

# Natively unfolded domains in endocytosis: hooks, lines and linkers

Timothy R. Dafforn<sup>1</sup> & Corinne J.I. Smith<sup>2\*</sup>

<sup>1</sup>University of Birmingham, Birmingham, UK, and <sup>2</sup>University of Warwick, Coventry, UK

It is commonly assumed that a protein must adopt a tertiary structure to achieve its active native state and that regions of a protein that are devoid of  $\alpha$ -helix or  $\beta$ -sheet structures are functionally inert. Although extended proline-rich regions are recognized as presenting binding motifs to, for example, Src homology 2 (SH2) and SH3 domains, the idea persists that natively unfolded regions in functional proteins are simply 'spacers' between the folded domains. Such a view has been challenged in recent years and the importance of natively unfolded proteins in biology is now being recognized. In this review, we highlight the role of natively unfolded domains in the field of endocytosis, and show that some important endocytic proteins lack a traditionally folded structure and harbour important binding motifs in their unstructured linker regions.

Keywords: adaptor proteins; clathrin; endocytosis; random coil; unstructured

EMBO reports (2004) 5, 1046–1052. doi:10.1038/sj.embor.7400276

## Introduction

The phenomenon of endocytosis, in which a cell membrane invaginates to enclose the desired cargo, pinches off and then fuses to the membrane of a target organelle, was first described by Roth & Porter (1964) during their work on yolk uptake by mosquito oocytes. Since then, the perceived role of endocytosis as a mechanism for nutrient uptake by the cell has broadened extensively to encompass protein transport, receptor downregulation, synaptic-vesicle recycling and, more recently, cell signalling and differentiation (Schmid, 1997; Jarousse & Kelly, 2001; Piddini *et al*, 2003). As our understanding of endocytosis has increased, so has the number of proteins that are known to be involved. Consequently, what seemed in 1964 to be a straightforward mechanism for nutrient uptake has become the hub of a network of vital cellular processes that are controlled by a wide range of factors.

## Many endocytic proteins have several binding partners

During endocytosis, specific proteins assemble on the membrane as it invaginates, many of which are included in the protein coat that surrounds the newly formed vesicle (reviewed in Brodsky *et al*, 2001). The components of this coat were among the first proteins to be identified as having a role in endocytosis; these include clathrin and the adaptor protein (AP2) complex, which were isolated from clathrin-coated vesicles (Pearse & Robinson, 1984; Ahle *et al*, 1988; Heuser & Keen, 1988). AP2 was subsequently discovered to be instrumental in selecting the contents of the vesicle, partly by binding specifically to motifs that are present in receptor tails on the cytoplasmic face of the plasma membrane (Ohno *et al*, 1995). The AP2 complex therefore earned its name by providing a link between clathrin on the outside of the coat, and the receptors and their cargo on the outside of the cell. AP2 consists of two large subunits ( $\alpha$ - and  $\beta$ 2-adaptin), a medium subunit ( $\mu$ 2) and a small subunit ( $\sigma$ 2). The large subunits are both composed of a 70–75-kDa trunk domain that is connected to an ~30-kDa 'ear' domain by an ~100-residue protease-sensitive 'hinge' region. Clathrin and AP2 have many binding partners; most of these have a role in endocytosis, but not all are part of the clathrin coat (reviewed in Lafer, 2002; Traub, 2003).

## Extensively unstructured regions in endocytic proteins

AP180, auxilin and epsin are all necessary for the formation of the clathrin coat, and circular dichroism data have shown that large proportions of their protein sequences are devoid of  $\alpha$ -helices or  $\beta$ -sheets (Kalthoff *et al*, 2002; Scheele *et al*, 2003). Although these regions are active, they seem to be unstructured. It is not unusual to find unstructured regions linking individual domains; the AP2 proteolytically sensitive hinge region is a well-studied example in endocytosis. However, the fact that so many of these functional proteins are unfolded challenges the commonly held perception of how a 'normal' protein should behave. This raises the question of whether other endocytic proteins also have a high proportion of unfolded structure. A strong indication of the potential secondary structure of a protein can be gained through secondary-structure-prediction analysis. Table 1 and Fig 1 show the results of a secondary-structure analysis of several endocytic proteins, which were calculated from the consensus of eight secondary-structure-prediction algorithms. The choice of proteins in this example was

<sup>1</sup>Department of Biosciences, University of Birmingham, Birmingham B15 2TT, UK

<sup>2</sup>Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

\*Corresponding author. Tel: +44 24 76 52 2 461; Fax: +44 024 76 52 3 568;

E-mail: corinne.smith@warwick.ac.uk

Submitted 14 June 2004; accepted 15 September 2004

**Table 1** | Predicted secondary structure type of some endocytic proteins

	$\alpha$ -helix	Extended $\beta$ -sheet	Random	Ambiguous
Clathrin HC	52.8*	8.18	32.96	6.03
Clathrin LC	49.38*	3.29	44.44	2.88
AP2-( $\alpha$ 2)	48.62*	5.94	39.51	5.94
AP2-( $\beta$ 2)	48.95*	6.84	38.53	5.68
AP2-( $\mu$ 2)	19.54	20.92	51.03*	8.51
AP2-( $\sigma$ 2)	59.15*	12.68	21.83	6.34
$\beta$ -arrestin	15.83	24.46	52.76*	6.95
AP180	32.08	3.97	60.30*	3.64
CALM	36.20	7.36	51.79*	4.75
Auxilin	19.78	13.30	63.30*	3.36
GAK	27.97	11.19	56.55*	4.29
Eps 15	49.92*	4.17	42.74	3.17
Epsin	33.39	2.18	61.89*	2.54
Endophilin	53.41*	2.84	40.06	3.69
Amphiphysin	35.97	7.91	51.51*	4.60
Dynamin	45.48*	7.64	42.54	4.35
Synaptojanin	23.57	16.01	56.35*	4.07
Synaptotagmin	26.13	18.76	48.46*	6.65
AAK1	25.84	11.36	59.56*	3.24
Stonin 1	21.50	15.37	58.91*	4.22
Stonin 2	16.69	15.14	62.43*	5.75
HIP1	67.23*	2.9	25.37	4.49
HIPR	69.38*	3.37	23.31	3.93
ARH	40.91	10.71	46.43*	1.95
HRS	38.29	4.61	54.61*	2.50
Dab2	13.38	10.65	71.95*	4.03
Numb	27.91	8.70	59.77*	3.61
GGA	15.07	16.71	65.21*	3.01

\*The most prevalent structure type for each protein. This table presents the percentages of each type of secondary structure found in a selection of proteins that are involved in clathrin-mediated endocytosis. The data were calculated from secondary-structure predictions performed using the NPS@: Network Protein Sequence Analysis web server (Combet *et al.*, 2000), which computes a consensus of several different secondary-structure-prediction algorithms. AAK1, adaptor-associated kinase 1; AP, adaptor protein; ARH, autosomal recessive hypercholesterolaemia; CALM, clathrin-assembly lymphoid myeloid leukaemia protein; clathrin HC/LC, clathrin heavy chain/light chain; Dab2, disabled 2; Eps15, epidermal growth factor-receptor substrate 15; GAK, cyclin-G-associated kinase; GGA, Golgi-associated  $\gamma$ -adaptin homology ARF-binding protein; HIP1, huntingtin-interacting protein 1; HIPR, huntingtin-interacting protein 1-related; HRS, hepatocyte growth factor-regulated tyrosine kinase substrate.

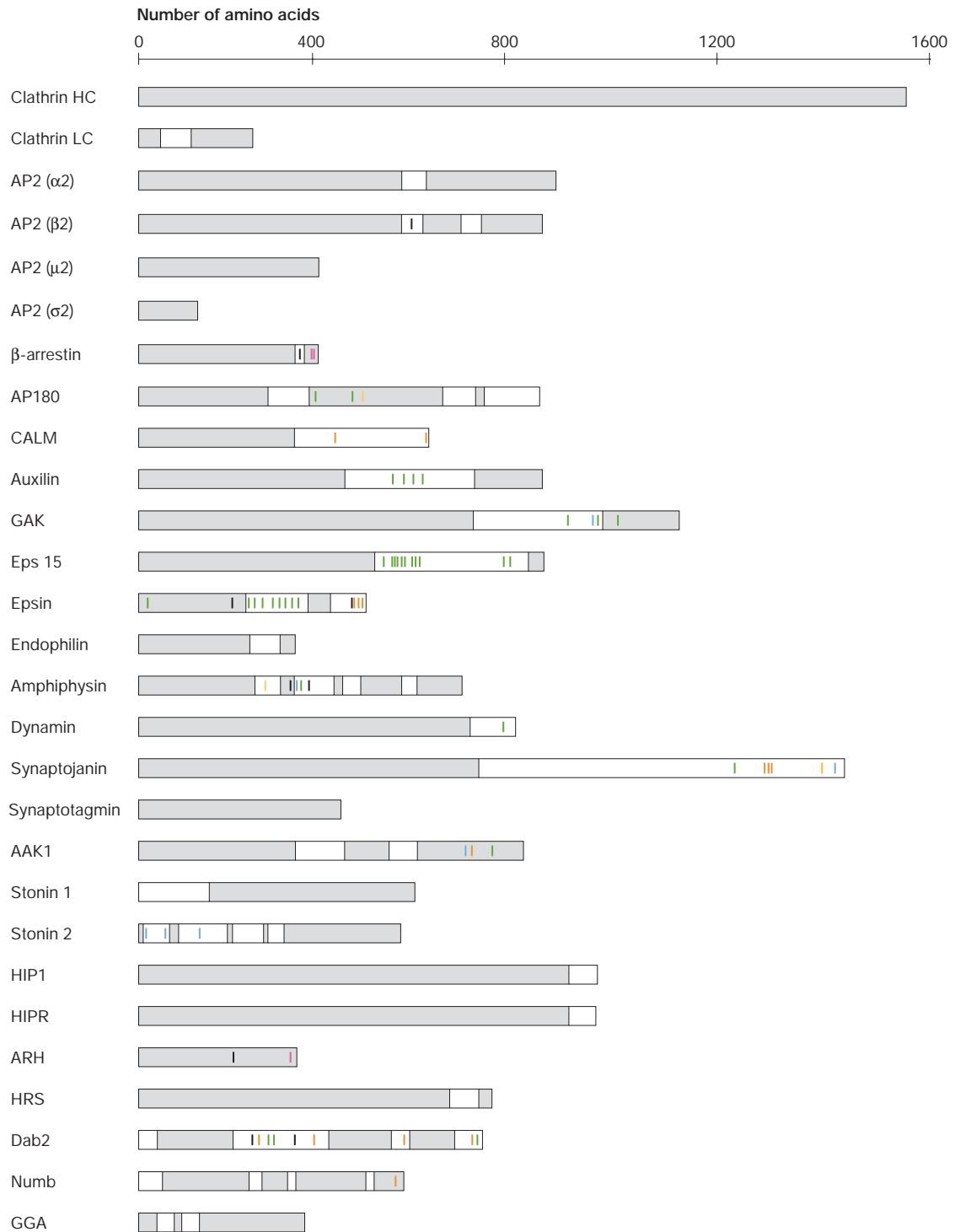
confined to those that are involved in clathrin-mediated endocytosis, and primarily to those that bind to clathrin or AP2. However, other proteins that are involved in transport and sorting might have similar properties. Figure 1 shows only long stretches of predicted random coil and illustrates how these are distributed throughout the protein sequences. Any continuous sequence of >40 residues that is predicted to be devoid of recognizable secondary structure has been classed as unfolded and shaded white; therefore, only extensive stretches of the predicted unfolded domains are shown. This is in contrast to Table 1, which presents the total predicted random-coil content. It should be noted that the precise definition of the term 'unfolded' is subject to debate and is not necessarily confined to completely random states.

Here, we highlight the range of structures that cannot be described as  $\alpha$ -helix or  $\beta$ -sheet, and which are extended in character and are solvent exposed. Figure 1 shows that many of the proteins that were analysed have long patches of continuous predicted random coil, which is contrary to what would be expected for a fully folded globular protein.

### The importance of natively unfolded proteins

In recent years, increasing attention has been focused on the fact that many proteins are neither folded nor do they have substantial unfolded domains under physiological conditions (Wright & Dyson, 1999; Tompa, 2002). These proteins have been described as 'natively unfolded' or 'intrinsically unstructured'. In fact, it has been predicted that >30% of all eukaryotic proteins are of this type (Dunker *et al.*, 2000; Ward *et al.*, 2004). In his discussion of intrinsically unstructured proteins (IUPs), Tompa (2002) summarized the recent findings and proposed that the unstructured state deserves recognition as a separate functional and structural entity. Could the endocytic proteins under discussion here be members of this class? To answer this question, experimental data from all of the proteins, rather than the few examples that are available, are required. However, it might be instructive to compare the secondary-structure-prediction data from Fig 1 with what is known for IUPs.

Figure 2 shows the amino-acid content of the endocytic proteins that were analysed as described for Fig 1, compared with those of globular proteins from the protein database, the IUPs analysed by Tompa (2002) and the amino-acid composition of collagen—a well-studied example of a protein that has an extended structure in its native form. IUPs are characterized by high levels of proline and charged amino acids compared with those of folded domains. The high prevalence of charged residues in these unstructured regions is thought to promote extension of the peptide chain, which suppresses the formation of a collapsed globular form (Dunker *et al.*, 2001; Tompa, 2002). The amino-acid composition of predicted unfolded regions from the analysis of endocytic proteins shows an elevated frequency of proline residues (20%) compared with globular proteins; this is almost twice the frequency that was reported for the IUPs analysed by Tompa (2002). Proline-rich regions in endocytic proteins have been observed surrounding the Src homology 3 (SH3)-binding motifs of dynamin and synaptojanin (Gout *et al.*, 1993; Wigge *et al.*, 1997; Cestra *et al.*, 1999), and also in, for example, adaptor-associated kinase 1 (AAK1; Conner & Schmid, 2002). In addition, the unfolded regions in the endocytic proteins show a significant increase in the number of smaller amino acids (serine, 12%; threonine, 8%; glycine, 10%; and alanine, 10%) compared with those observed for folded proteins and IUPs, and charged amino acids are also underrepresented. This composition is similar to that predicted for the collagen and collagen-like family of proteins (Fig 2). It is possible that the high proportion of serine and threonine residues might point to the importance of phosphorylation sites in regions of protein disorder. Iakoucheva *et al.* (2004) noted that disorder-promoting residues tend to surround phosphorylation sites in the SWISS-PROT protein database and there is no doubt that phosphorylation is an important control mechanism in endocytosis (reviewed in Korolchuk & Banting, 2003). Hence, there seem to be both similarities and differences between IUPs, folded proteins and the class of endocytic proteins



**Fig 1** | Illustration of the locations of areas of predicted random coil in a selection of endocytic proteins. The consensus of eight secondary-structure-prediction algorithms indicates large stretches of unfolded chain (white) in the sequences of a selection of proteins that are involved in clathrin-mediated endocytosis. A region of extended chain is differentiated from shorter loop regions by defining it as a region of >40 residues in length that is consistently predicted as unstructured. The motifs are represented as coloured lines as follows: DPF or DPW, green; NPF, orange; WXX[FW]X[DE], cyan; FXDXF motifs, yellow;  $\beta$ 2-adaptin ear-binding motifs, purple; and clathrin-binding motifs, black. Although this diagram is not drawn to scale, it is a reasonable representation of the relative positions of the features marked. The abbreviations are defined in Table 1.

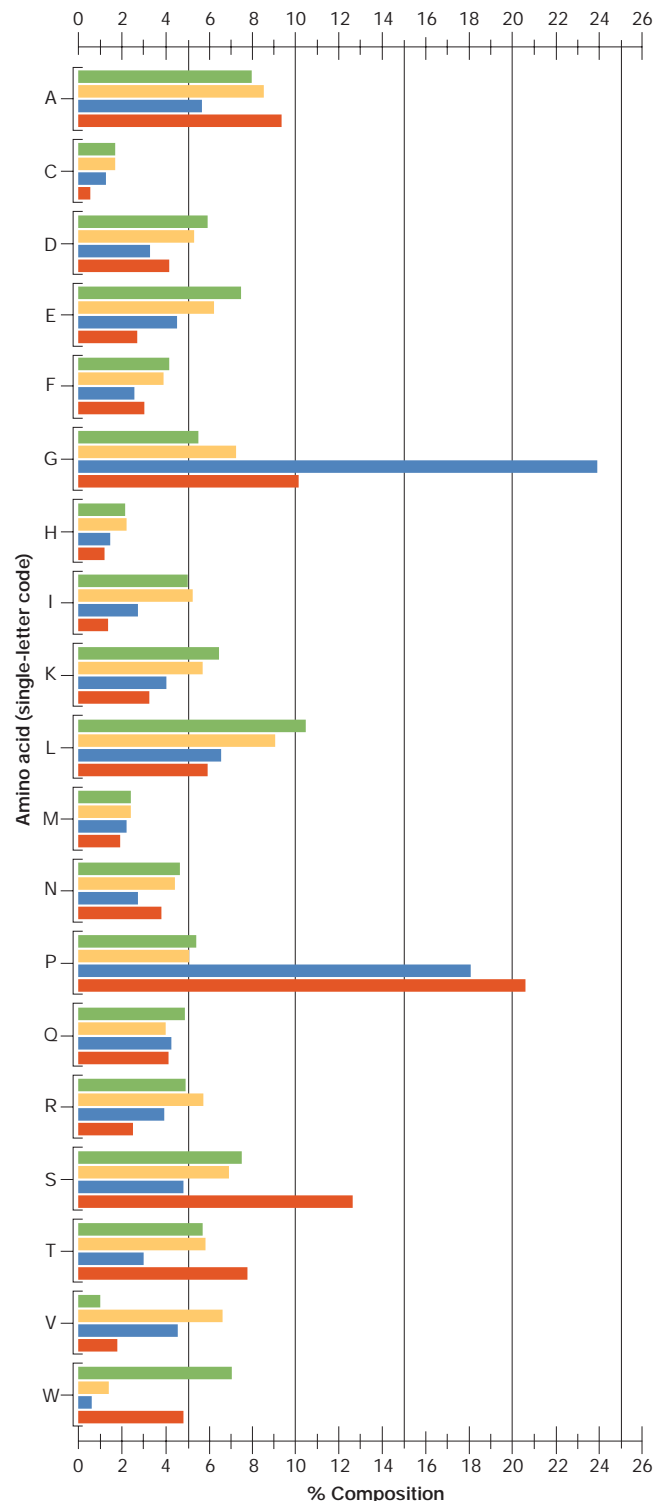
that is under discussion; indeed, these proteins might actually represent a specific subset of IUPs.

### Low-affinity binding motifs

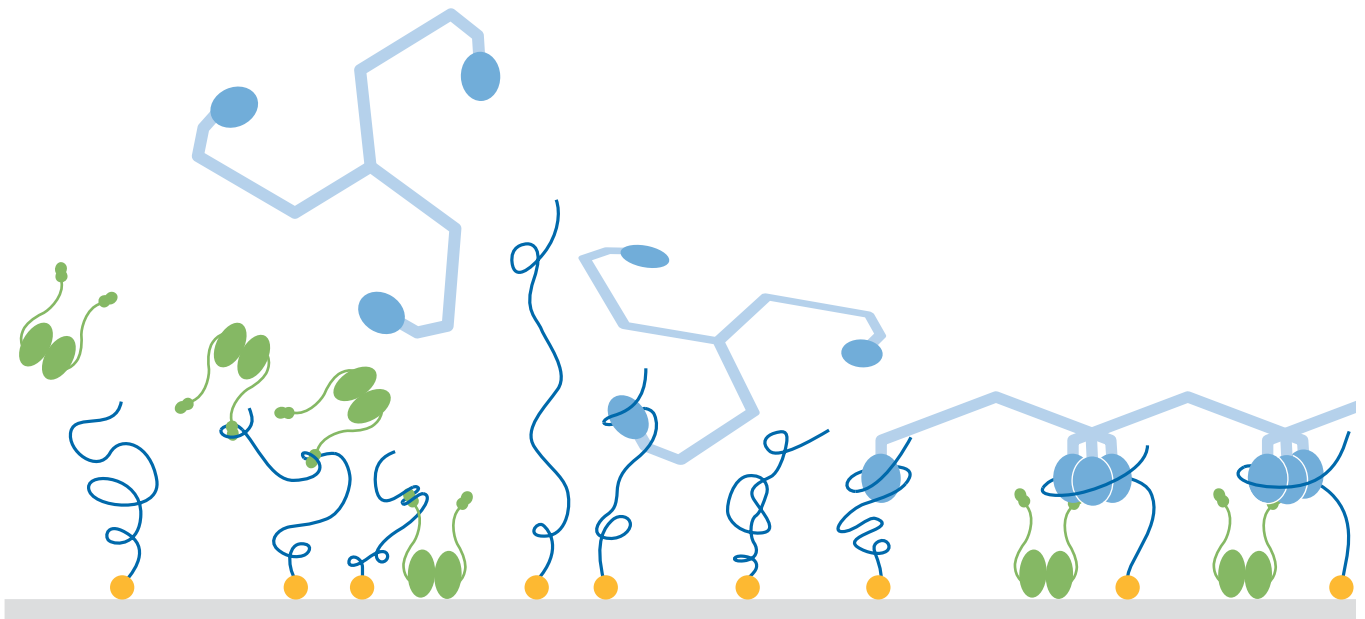
For endocytosis to occur, a wide range of connections between proteins must be coordinated accurately and rapidly, and this is frequently achieved by the use of low-affinity protein–peptide interactions. The ‘clathrin box’ binding motif that is found in the proteolytically sensitive hinge region of AP2  $\beta$ 2-adaptin is one such example (Shih *et al*, 1995). This motif, which is loosely described by the sequence LLNLD, has been identified in many other clathrin-binding proteins and has been shown, in some cases, to bind to the clathrin amino-terminal domain (Dell’Angelica, 2001). A second example of the use of such short peptide-binding motifs is found in the binding of the  $\mu$ 2-subunit of AP2 to YXX $\Phi$  motifs (in which  $\Phi$  is a bulky hydrophobic amino acid) on the cytoplasmic tail of receptors that are destined for recruitment to the vesicle (Ohno *et al*, 1995). In addition, tripeptide NPF motifs mediate protein binding to the Eps homology (EH) domains of the epidermal-growth-factor-receptor substrate 15 (Eps15; Salcini *et al*, 1997), and proline-rich sequences in dynamin and synaptojanin are used to bind to the SH3 domain of amphiphysin (David *et al*, 1996; McPherson *et al*, 1996; Grabs *et al*, 1997). In the case of AP2, the tripeptide DPF/W, the FXDXF motif and, most recently, the WXX[FW]X[DE] motif have all been identified as binding to the AP2  $\alpha$ -adaptin ear domain (Brett *et al*, 2002; Mishra *et al*, 2004).

The degree to which these motifs have been characterized in each of the proteins in which they occur varies greatly. However, the binding constants that have been determined are generally in the range of 5–500  $\mu$ M (de Beer *et al*, 1998; Miele *et al*, 2004), which is relatively weak compared with, for example, enzyme–substrate or antibody–antigen interactions. The finding that such weak protein–peptide interactions can determine the complex and specific needs of endocytosis challenges the theory that a specific interaction must be of high affinity to be physiologically significant. Indeed, examination of the functions of unfolded regions in other proteins has highlighted their advantages in providing high specificity as well as low-affinity binding sites (Dunker *et al*, 2002).

Structural data are available that show how peptides relating to most of these motifs bind. In two highly illuminating studies, both the clathrin terminal domain and the AP2  $\alpha$ -adaptin ear domain have been shown to contain more than one motif-binding site (Miele *et al*, 2004; Mishra *et al*, 2004). The work of Brett *et al* (2002) and Mishra *et al* (2004) has revealed the nature of the binding sites for three motifs on the AP2  $\alpha$ -adaptin ear: DP[FW], FXDXF and WXX[FW]X[DE]. These studies show that DP[FW] and FXDXF have overlapping sites, whereas the site for WXX[FW]X[DE] is located on a separate domain. It seems that AP2 uses a mixture of competitive and non-competitive binding to these motifs, thereby allowing important regulatory proteins, such as synaptojanin, to gain access through the WXX[FW]X[DE] site, whilst also ensuring the appropriate selection of cargo-carrying proteins, such as AP180, through the DP[FW] and FXDXF sites. Mishra *et al* (2004) have highlighted the physiological importance of competitive binding in the case of epsin, by showing a correlation between its relatively high affinity for the  $\alpha$ -adaptin ear domain, its retention during vesicle formation and its inclusion in the clathrin-coated vesicle.



**Fig 2** | The percentage amino-acid composition of the predicted unstructured and structured regions of the group of endocytic proteins detailed in Fig 1. The regions that are predicted to be structured are shown in green and those expected to be unstructured are shown in red. The average amino-acid composition of globular proteins is shown in yellow (Creighton, 1993), and the amino-acid composition of human type I collagen is shown in blue.



**Fig 3** | Illustration of the possible role of natively unfolded domains of epsin, adaptor protein 180 (AP180) and AP2 during clathrin-coat formation. After binding to the membrane (Ford *et al*, 2001, 2002), the long flexible domains of AP180 and epsin have the potential to bind to several sites on AP2 and clathrin. They might also bind to many clathrin or AP2 molecules. In combination with the self-assembly properties of clathrin triskelions, this would result in a highly cooperative assembly mechanism (Takei & Haucke, 2001). This figure is adapted from Kalthoff *et al* (2002) and is reproduced with permission from the American Society for Biochemistry and Molecular Biology.

**The ‘fishing’ theory of endocytosis**

It has recently been noted (Brett *et al*, 2002; Evans & Owen, 2002; Kalthoff *et al*, 2002) that unfolded linker regions or larger domains are important in many endocytic proteins. Interestingly, the short peptide-binding motifs described above seem to be coincident with such regions. This is the case for the clathrin-box motif in the  $\beta$ 2-adaptin hinge region of AP2, and, certainly, most of the DPF and NPF motifs that are found in AP180 and epsin fall in unfolded domains. This pattern is being observed with increasing regularity for disordered regions in other proteins. In fact, it has recently been proposed that binding sites on extended regions have the added advantage of allowing several targets to bind with relatively low affinity (Dunker *et al*, 2001). Studies of these regions in endocytic proteins have led Kalthoff *et al* (2002) to propose that AP180 and epsin mainly consist of a long flexible polypeptide that is attached to the membrane by a folded lipid-binding domain. The long flexible chain can access binding partners from a relatively large volume of the cytosol, whilst remaining attached to the membrane. Figure 3, which is based on ideas proposed by Kalthoff *et al* (2002), illustrates how such interactions might be used during clathrin-coat formation. Evans & Owen (2002) have also taken up the theme and have described this concept as ‘protein fishing’, in which an unfolded protein ‘string’ that contains several binding sites is able to ‘hook’ binding partners. In addition to the proteins analysed by Kalthoff *et al* (2002), Evans & Owen (2002) refer to the presence of the clathrin-box-binding motif on the unstructured  $\beta$ 2-adaptin hinge region and to the dileucine motif on the hinge of the Golgi-associated  $\gamma$ -adaptin homology ARF-binding (GGA) protein.

**Do unfolded domains contain several binding motifs?**

The idea of protein fishing through many linear low-affinity interaction motifs on an unstructured protein fishing line has been discussed here in relation to four main examples: AP180, epsin, AP2  $\beta$ 2-adaptin and GGA. However, it is also of interest to determine how widespread this phenomenon is for endocytic proteins. Figure 1 displays a range of proteins that have been analysed in terms of the location of extensive stretches of predicted random coil matched with the occurrence of some of the binding motifs. The white shading indicates continuous stretches of predicted random coil. Segments of random coil of less than 40 residues have been excluded; therefore, to be labelled as such, the stretch of random coil must already be sufficiently long. On the basis of the incidence of linear peptide-binding motifs, 84% of these cases fall in the regions of predicted random coil. Some deviations from the theory—for example, in the case of AP180—can be resolved by examination of the entire predicted protein structure. In AP180, the motifs are, in fact, on predicted loop regions that do not qualify as random coils in this study, but which have a notably unfolded character. Other cases can also be rationalized by the presence of a predicted random coil that fails the stringent test applied here. By contrast, in the autosomal recessive hypercholesterolaemia (ARH) protein, the verified clathrin box falls in a strongly predicted helix. Overall, however, it seems that these unstructured regions are generally used as a means to present binding sites to potential binding partners.



### Biological advantages of an unstructured region

*Accessibility of the binding motif.* Owing to the extended nature of these linear peptide-binding motifs in the context of unfolded polypeptides, they are possibly the most accessible of all binding sites. For example, an extended polypeptide of 50 residues would have a potential span of 175 Å. A single protein with a significant unfolded domain could connect with other proteins or the membrane across relatively large distances. It is therefore possible for such a protein to present several binding sites to a relatively wide area, and to thereby improve the efficiency with which proteins are assembled or become recruited to the membrane. Mechanisms involving many binding sites in the form of sequential arrays of folded domains are not uncommon in biology and occur, for example, in the extracellular matrix.

*Space efficiency.* The second advantage, conversely, is in saving space. Long, thin unfolded polypeptides on a membrane pack laterally into a much smaller 'footprint' than their folded counterparts. In situations in which several proteins require access to a single binding site or where there are adjacent binding sites, as in the case of the AP2  $\alpha$ -adaptin ear domain, the small diameter of a linear motif on an unfolded polypeptide chain should allow all of the binding partners to remain close together. This would facilitate the efficient exchange of binding partners after each binding interaction has fulfilled its purpose.

*Cooperativity.* The combination of tethering to, for example, the membrane and several sites of interaction covering long distances increases the potential for cooperative binding. Physical tethering of a protein close to a weak binding site will create a high local protein concentration, thereby allowing the low-affinity binding site to become 'active'. This must be especially true when several copies of a binding motif are present along a length of polypeptide chain. Conversely, once the proteins in question are untethered, weak interactions will not prevent the release of the protein. Therefore, such a mechanism should facilitate efficient assembly and disassembly.

### Final comments

The many binding partners of clathrin and AP2 raise issues of coordination and control in the cell. The work of Miele *et al* (2004) and Mishra *et al* (2004) indicates that, in the cases of the clathrin terminal domain and the AP2  $\alpha$ -adaptin ear, different motifs on the same accessory protein might bind to two separate sites on clathrin or AP2. Such a model of binding makes excellent use of the flexible properties of an extended polypeptide chain. These extended regions have the potential to bind more than one molecule of clathrin or AP2 per chain, thereby increasing the efficiency of the recruitment and assembly of these proteins during coat formation. Finally, as the coat assembles and detaches, the small diameter of these extended domains might facilitate optimal packing of the many proteins that are included within the confines of the clathrin coat. Although much remains to be understood, it is clear that the flexibility, length and small diameter of natively unfolded polypeptides are of prime importance during endocytosis.

### ACKNOWLEDGEMENTS

The authors thank B. Pearce for critical reading of the manuscript. T.R.D. is a Medical Research Council (MRC) Fellow and C.J.S thanks the MRC for support.

### REFERENCES

- Ahle S, Mann A, Eichelsbacher U, Ungewickell E (1988) Structural relationships between clathrin assembly proteins from the Golgi and the plasma-membrane. *EMBO J* **7**: 919–929
- Brett TJ, Traub LM, Fremont DH (2002) Accessory protein recruitment motifs in clathrin-mediated endocytosis. *Structure (Camb)* **10**: 797–809
- Brodsky FM, Chen CY, Knuehl C, Towler MC, Wakeham DE (2001) Biological basket weaving: formation and function of clathrin-coated vesicles. *Annu Rev Cell Dev Biol* **17**: 517–568
- Cestra G, Castagnoli L, Dente L, Minenkova O, Petrelli A, Migone N, Hoffmuller U, Schneider-Mergener J, Cesareni G (1999) The SH3 domains of endophilin and amphiphysin bind to the proline-rich region of synaptojanin 1 at distinct sites that display an unconventional binding specificity. *J Biol Chem* **274**: 32001–32007
- Combet C, Blanchet C, Geourjon C, Delaage G (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci* **25**: 147–150
- Conner SD, Schmid SL (2002) Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J Cell Biol* **156**: 921–929
- Creighton TE (1993) *Proteins: Structures and Molecular Properties*. New York, USA: Freeman
- David C, McPherson PS, Mundigl O, de Camilli P (1996) A role of amphiphysin in synaptic vesicle endocytosis suggested by its binding to dynamin in nerve terminals. *Proc Natl Acad Sci USA* **93**: 331–335
- de Beer T, Carter RE, Lobel-Rice KE, Sorkin A, Overduin M (1998) Structure and Asn-Pro-Phe binding pocket of the Eps15 homology domain. *Science* **281**: 1357–1360
- Dell'Angelica EC (2001) Clathrin-binding proteins: got a motif? Join the network! *Trends Cell Biol* **11**: 315–318
- Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ (2000) Intrinsic protein disorder in complete genomes. *Genome Inform Ser Workshop Genome Inform* **11**: 161–171
- Dunker AK *et al* (2001) Intrinsically disordered protein. *J Mol Graph Model* **19**: 26–59
- Dunker AK *et al* (2002) Intrinsic disorder and protein function. *Biochemistry* **41**: 6573–6582
- Evans PR, Owen DJ (2002) Endocytosis and vesicle trafficking. *Curr Opin Struct Biol* **12**: 814–821
- Ford MG, Pearce BM, Higgins MK, Vallis Y, Owen DJ, Gibson A, Hopkins CR, Evans PR, McMahon HT (2001) Simultaneous binding of PtdIns(4,5)P<sub>2</sub> and clathrin by AP180 in the nucleation of clathrin lattices on membranes. *Science* **291**: 1051–1055
- Ford MG, Mills IG, Peter BJ, Vallis Y, Praefcke GJ, Evans PR, McMahon HT (2002) Curvature of clathrin-coated pits driven by epsin. *Nature* **419**: 361–366
- Gout I *et al* (1993) The GTPase dynamin binds to and is activated by a subset of SH3 domains. *Cell* **75**: 25–36
- Grabs D, Slepnev VI, Songyang Z, David C, Lynch M, Cantley LC, De Camilli P (1997) The SH3 domain of amphiphysin binds the proline-rich domain of dynamin at a single site that defines a new SH3 binding consensus sequence. *J Biol Chem* **272**: 13419–13425
- Heuser JE, Keen J (1988) Deep-etch visualization of proteins involved in clathrin assembly. *J Cell Biol* **107**: 877–886
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* **32**: 1037–1049
- Jarousse N, Kelly RB (2001) Endocytotic mechanisms in synapses. *Curr Opin Cell Biol* **13**: 461–469
- Kalthoff C, Alves J, Urbanke C, Knorr R, Ungewickell EJ (2002) Unusual structural organization of the endocytic proteins AP180 and epsin 1. *J Biol Chem* **277**: 8209–8216
- Korolchuk V, Banting G (2003) Kinases in clathrin-mediated endocytosis. *Biochem Soc Trans* **31**: 857–860
- Lafer EM (2002) Clathrin-protein interactions. *Traffic* **3**: 513–520
- McPherson PS, Garcia EP, Slepnev VI, David C, Zhang X, Grabs D, Sossin WS, Bauerfeind R, Nemoto Y, De Camilli P (1996) A presynaptic inositol-5-phosphatase. *Nature* **379**: 353–357
- Miele AE, Watson PJ, Evans PR, Traub LM, Owen DJ (2004) Two distinct interaction motifs in amphiphysin bind two independent sites on the clathrin terminal domain  $\beta$ -propeller. *Nat Struct Mol Biol* **11**: 242–248

Mishra SK, Hawryluk MJ, Brett TJ, Keyel PA, Dupin AL, Jha A, Heuser JE, Fremont DH, Traub LM (2004) Dual-engagement regulation of protein interactions with the AP-2 adaptor  $\alpha$ -appendage. *J Biol Chem* 2 Aug [epub ahead of print]

Ohno H, Stewart J, Fournier MC, Bosshart H, Rhee I, Miyatake S, Saito T, Gallusser A, Kirchhausen T, Bonifacino JS (1995) Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science* **269**: 1872–1875

Pearse BMF, Robinson MS (1984) Purification and properties of 100-Kd proteins from coated vesicles and their reconstitution with clathrin. *EMBO J* **3**: 1951–1957

Piddini E, Vincent JP, Murthy VN, De Camilli P, Jarrow N, Kelly RB (2003) Modulation of developmental signals by endocytosis: different means and many ends. *Curr Opin Cell Biol* **15**: 474–481

Roth TF, Porter KR (1964) Yolk protein uptake in the oocyte of the mosquito *Aedes aegypti*. *J Cell Biol* **20**: 313–332

Salcini AE, Confalonieri S, Doria M, Santolini E, Tassi E, Minenkova O, Cesareni G, Pellicci PG, Di Fiore PP (1997) Binding specificity and *in vivo* targets of the EH domain, a novel protein–protein interaction module. *Genes Dev* **11**: 2239–2249

Scheele U, Alves J, Frank R, Duwel M, Kalthoff C, Ungewickell EJ (2003) Molecular and functional characterization of clathrin and AP-2 binding determinants within a disordered domain of auxilin. *J Biol Chem* **278**: 25357–25368

Schmid SL (1997) Clathrin-coated vesicle formation and protein sorting: an integrated process. *Annu Rev Biochem* **66**: 511–548

Shih W, Gallusser A, Kirchhausen T (1995) A clathrin-binding site in the hinge of the  $\beta$ 2 chain of mammalian AP-2 complexes. *J Biol Chem* **270**: 31083–31090

Takei K, Haucke V (2001) Clathrin-mediated endocytosis: membrane factors pull the trigger. *Trends Cell Biol* **11**: 385–391

Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27**: 527–533

Traub LM (2003) Sorting it out: AP-2 and alternate clathrin adaptors in endocytic cargo selection. *J Cell Biol* **163**: 203–208

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* **337**: 635–645

Wigge P, Vallis Y, McMahon HT (1997) Inhibition of receptor-mediated endocytosis by the amphiphysin SH3 domain. *Curr Biol* **7**: 554–560

Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *J Mol Biol* **293**: 321–331



**Timothy R. Dafforn**



**Corinne J. I. Smith**