



**HAL**  
open science

## Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis

Nurzian Ismail, Samuel G Crawshaw, Benedict C.S. Cross, Anna C Haagsma,  
Stephen High

### ► To cite this version:

Nurzian Ismail, Samuel G Crawshaw, Benedict C.S. Cross, Anna C Haagsma, Stephen High. Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis. *Biochemical Journal*, 2008, 411 (3), pp.495-506. 10.1042/BJ20071597 . hal-00478923

**HAL Id: hal-00478923**

**<https://hal.science/hal-00478923>**

Submitted on 30 Apr 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis

Nurzian Ismail<sup>1</sup>, Samuel G. Crawshaw, Benedict C. S. Cross, Anna C. Haagsma<sup>2</sup> and Stephen High\*

Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK.

<sup>1</sup>Current address: Temasek Life Sciences Laboratory, Cell Stress and Homeostasis Group, 1 Research Link, NUS, Singapore 117604.

<sup>2</sup>Current address: Department of Structural Biology, Institute of Molecular Cell Biology, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.

\* Author for correspondence: TEL. +44 (0)161 275 5070; FAX. +44 (0)161 275 5082; Email: stephen.high@manchester.ac.uk

**Short (page heading) title:** TM segment delay at the ER translocon

**Keywords:** Independent folding domains/endoplasmic reticulum/membrane protein assembly/PAT-10/cross-linking/Sec61

**Abbreviations used:** BMH, bismaleimido-hexane; EndoH, endoglycosidase H; ER, endoplasmic reticulum; HA, haemagglutinin; OP, bovine opsin; OPN, N-terminal region of opsin; OPN/5-7, OPN fused to TM segments 5-7; OPN/TM5,6,3, OPN/5-7 with TM3 in place of TM7; OPN/TM3, OPN fused to opsin TM3; PAT-10, protein associated with the ER translocon of 10 kDa; PPL, preprolactin; TRAM, translocating chain-associated membrane protein; TM, transmembrane.

## SYNOPSIS

A site-specific cross-linking approach was used to study the integration of transmembrane (TM) segments 4-7 of the polytopic membrane protein, opsin, at the endoplasmic reticulum (ER). We find that whilst TM4 exits the ER translocon rapidly, TM segments 5, 6 and 7 are all retained at the translocon until opsin biosynthesis is terminated. Furthermore, although artificial extension of the nascent chain is not sufficient to release the C-terminal region of opsin from the translocon, substitution of the native TM segment 7 with a more hydrophobic TM segment results in its rapid lateral exit into the lipid bilayer. We conclude that the intrinsic properties of a TM segment determine the timing of its membrane integration rather than its relative location within the polypeptide chain. A pronounced and prolonged association of opsin TM5 with the translocon associated component PAT-10 was also observed, suggesting that PAT-10 may facilitate the assembly of distinct opsin sub-domains during membrane integration. Our data strongly support a model in which the ER translocon co-ordinates the integration of selected TM segments in response to the specific requirements of the precursor being synthesised.

## INTRODUCTION

A crucial step in the biogenesis of integral membrane proteins is their incorporation into the hydrophobic environment of the lipid bilayer [1, 2]. In eukaryotic cells this event frequently occurs co-translationally at the Sec61 translocon of the endoplasmic reticulum (ER), a site that also mediates the complete translocation of soluble proteins into the ER lumen where conditions are optimised for folding and maturation [1-3].

The core of the ER translocon comprises a heterotrimer of Sec61 $\alpha$ ,  $\beta$ , and  $\gamma$  subunits [4]. High resolution structural studies of a related archeal translocon indicate that protein translocation may occur through a narrow pore formed by a single Sec61 heterotrimer [5]. In this model, a translocation pore of ~5-8 Å in diameter would be gated by a subdomain of the Sec61 $\alpha$  subunit [5]. Whilst deletion of this putative 'plug' in both the yeast and bacterial homologues causes only minor defects in co-translational translocation [6, 7], it is suggested that replacement plug structures may be formed in these deletion mutants [8]. A structural model of the eubacterial SecYEG translocon, captured with a bound ribosome-nascent chain complex, suggests that two heterotrimers may act in concert to form a larger translocation pore through the membrane [9]. Likewise, cross-linking data from purified bacterial SecY complexes reconstituted into proteoliposomes supports a model in which oligomerisation of the core SecYEG heterotrimer facilitates the formation of an active translocation channel [10]. Thus, whilst the translocation pore itself may be formed by just one Sec61 heterotrimer, it is likely that multiple copies are required for translocation to occur [11-13]. Subsidiary translocon components, such as the translocating chain-associated membrane (TRAM) protein and a protein associated with the ER translocon of 10 kDa (PAT-10), are proposed to have specialised functions during protein biosynthesis at the ER although their roles are yet to be fully defined [14-17].

The hydrophobic TM segments present in membrane proteins appear to engage the ER translocon and then enter the lipid bilayer, presumably via a 'lateral gate' separating the aqueous translocation pore from the hydrophobic membrane phase [17, 18]. This 'partitioning' of TM segments into the lipid bilayer during translocation has been studied in some detail for simple, single-spanning membrane proteins [14, 15, 18, 19]. However, polytopic membrane protein integration presents additional problems, since multiple TM segments must be integrated and correctly assembled [1, 2]. Studies to date suggest that the integration of polytopic proteins is influenced by multiple factors, including: net hydrophobicity of a TM span [15, 20], TM segment interactions [21, 22] and the length of soluble loop domains connecting TM regions [23]. In the case of polytopic proteins such as bacterioopsin and aquaporin-4 (AQP4), studies to date suggest TM integration occurs sequentially as each new TM segment is synthesized by the ribosome [16, 24-27]. However, the fate of TM segments following initial engagement with the ER translocon can vary depending on the precursor, for example the synthesis of a new TM segment may not necessarily lead to the complete release of the preceding TM segment from the Sec61 complex. Hence, in the case of opsin, TM1 is 'parked' at a distinct Sec61 location after initial integration at the translocon core [16, 24] and a similar phenomenon is observed for multiple TM segments within AQP4 [26]. Thus, for limited number of polytopic membrane proteins studied to date, the integration process appears remarkably complex [17, 24, 26].

Both TM segments and soluble loop regions of complex polytopic membrane proteins must be correctly assembled to generate a functional conformation and these processes may be linked. Hence, in the case of the cystic fibrosis transmembrane conductance regulator (CFTR)

## TM segment delay at the ER translocon

protein, the correct folding of the first ATP binding domain present in a large intracellular loop is dependent on the assembly of preceding membrane spanning segments [28]. Similarly, intramembrane assembly of AQP1 is dependent on close intra- and intermolecular bonding between TM segments [22]. For AQP1, these interactions are dependent on highly polar residues within the TM segments that compromise hydrophobicity but appear necessary to achieve a correctly folded and functional protein structure.

We selected opsin as a model polytopic protein for the study of membrane integration at the ER [25], and established the use of site-specific chemical cross-linking as a valid approach to report the proteinaceous environment of individual TM segments [16]. To date we have analysed the behaviour of opsin TM segments 1-3 and carefully mapped their lateral exit from the ER translocon [16, 24]. These studies revealed that different TM segments associate with distinct translocon components [16] and can exit the translocon by distinct mechanisms [24], but did not address the integration of opsin TM segments 4-7. Here, we have specifically investigated the biosynthesis of the C-terminal region of opsin in order to complete our previous studies and provide a full molecular description of its integration [16, 24]. We show that whilst opsin TM4 is integrated into the lipid bilayer following the synthesis of TM5, TM segments 5, 6 and 7 all remain in close proximity to the ER translocon throughout opsin biosynthesis. Remarkably, extending the cytosolic C-terminal tail of opsin does not release these TM segments from the translocon, indicating that TMs 5-7 are not simply 'tethered' to the Sec61 complex. However, by replacing opsin TM7 with a different TM segment, rapid exit from the ER translocon can be achieved. The distinctive behaviour of TM7 at the ER translocon, together with the selective adduct formation between the opsin nascent chain and PAT-10, leads us to propose that the retention of the C-terminal TM segments of opsin at the Sec61 translocon reflects substrate specific assembly requirements. We incorporate this hypothesis into a unifying model for opsin biogenesis.

## MATERIALS AND METHODS

### Reagents

T7 RNA polymerase, transcription reagents, RNasin ribonuclease inhibitor, RNase A and nuclease-treated rabbit reticulocyte lysate were purchased from Promega (Southampton, UK). Easytag L-[<sup>35</sup>S]-methionine/cysteine was obtained from PerkinElmer Life Sciences (Stevenage, UK). The cross-linking reagent, bismaleimido-hexane (BMH), was purchased from Perbio Science (Chester, UK). All other chemical reagents were obtained from Sigma Chemical (Poole, UK) and Merck (Poole, UK). Reagents for cell culture were purchased from Invitrogen (Paisley, UK). The mouse monoclonal anti-haemagglutinin (HA) antibody was a gift from Dr. I. Hagan (Paterson Laboratories, Manchester, UK), and the rabbit polyclonal antibodies specific for the Sec61 $\alpha$  and Sec61 $\beta$  subunits were kindly provided by Prof. R. Zimmermann (University of Saarland, Germany) and Prof. B. Dobberstein (University of Heidelberg, Germany). The mouse monoclonal antibody specific for the N-terminal region of bovine opsin was purified from a hybridoma line originally supplied by Dr. P. Hargrave (University of Florida, USA) [29].

### Opsin constructs

Single or double cysteine residues were introduced using the QuikChange site-directed mutagenesis kit (Stratagene, Cambridge, UK) into a cysteine-null version of bovine opsin previously described [16]. The OPN/5-7 construct was generated by the deletion of residues 36 to 194 to give a polypeptide with opsin residues 1 to 35 fused with residues 195 to 348. For the extension of OPN/5-7 nascent chain beyond the last residue of opsin, the C-terminus of OPN/5-7 was PCR spliced with residues 31 to 220 of bovine prolactin and cloned into the PCR-SCRIPT vector (Stratagene) downstream of a T7 promoter. This was followed by the introduction of single cysteine residues at specified locations. The deletion of TM5 and TM6 to generate the OPN/TM7PPL construct was carried out using the ExSite mutagenesis kit according to the supplier's instructions (Stratagene). The replacement of TM7 with TM3 to generate the OPN/TM3PPL construct was completed using the QuikChange site directed mutagenesis kit (Stratagene). The replacement in TM7 with TM3 in the OPN/5-7-PPL construct to generate OPN/TM5,6,3PPL was also completed using the Quikchange kit.

### In vitro transcription and translation

mRNA was obtained by performing in vitro transcription with T7 RNA polymerase (Promega) as described by the manufacturer and purified using an RNeasy Mini Kit (Qiagen). In all experiments, digitonin-permeabilised cultured HT1080 fibroblasts (European Collection of Cell Cultures, Salisbury, UK) [30] were used as a source of ER membranes. In vitro translation of the mRNA chains was carried out in a rabbit reticulocyte lysate system (Promega) for 15 minutes at 30°C in the presence of [<sup>35</sup>S]methionine/cysteine (0.82  $\mu$ Ci/ $\mu$ l) and semi-intact HT1080 cells. Initiation of translation was subsequently inhibited by the addition of 0.1 mM aurintricarboxylic acid for 10 minutes at 30°C. Cycloheximide to 2 mM was added to terminate translation reactions and the samples were placed on ice for 5 minutes. Where nascent chains were released from the ribosome prior to cross-linking, samples were treated with 2 mM puromycin and 50 mM EDTA for 10 min at 30°C in place of the cycloheximide treatment. The membrane fraction was isolated by centrifugation at 16,000 g for 30 seconds and washed twice in KHM buffer (110 mM potassium acetate, 20 mM

HEPES, 2 mM magnesium acetate, pH 7.2). The final pellet was resuspended in KHM buffer for cross-linking reactions.

### **Cross-linking and immunoprecipitation**

For cross-linking, 1 mM BMH (final concentration) was added to the membrane fraction and the mixture incubated at 30°C for 10 minutes. Only DMSO was added for the solvent control. The cross-linking reaction was subsequently quenched with the addition of 2-mercaptoethanol to 5 mM final concentration and the samples treated with 250 µg/ml RNase A at 37°C for 5 minutes. Denaturing immunoprecipitations were then carried out as previously described [16].

### **Endoglycosidase H treatment**

Immunoprecipitated samples on Protein A beads were denatured in endoglycosidase H (Endo H) denaturing buffer (0.5% SDS, 1% β-mercaptoethanol) for 30 minutes at 37°C. 3 µl of 0.5M sodium citrate pH 5.5 buffer, 0.5 µl of 100 mM phenylmethyl sulphonyl fluoride, 5.5 µl of H<sub>2</sub>O and 500 U of Endo H were added and the samples were incubated at 37°C for another 30 minutes. The samples were solubilised in an equal volume of 2x concentration SDS sample buffer and resolved by SDS-PAGE.

### **Sample analysis**

Samples were solubilised in SDS-PAGE sample buffer at 37°C for 30 minutes and resolved on 14% Tris-glycine polyacrylamide gels. The resulting gels were fixed and dried before being exposed to a phosphorimager plate for visualization using Phosphor Imager FLA-3000 (Fuji, Kanagawa, Japan). We impose a strict criteria for authentic cross-linking adducts; the adduct should be clearly immunoprecipitated by both the α-opsin (Op) antiserum, which recognizes the N-terminus of the opsin chain, and the α-HA antiserum, which recognizes the HA epitope tag at the C-terminus. Identification of the cross-linking partner can then be achieved by immunoprecipitation of the adduct with antisera specific to translocon components. This allows true adducts to nascent chains with an intact N- and C-terminus to be distinguished from other products which may be formed with truncated opsin chains [24].

## RESULTS

Our previous studies of opsin biosynthesis have primarily focused on the molecular environment of an N-terminal region of opsin, comprising TM segments 1 to 3, during its integration into the ER membrane [16, 24, 25]. Here, we extend this work to examine the integration of the remainder of the opsin polypeptide incorporating TM segments 4 to 7. Site-specific cross-linking was employed to analyse the molecular environment of each TM segment using integration intermediates of increasing chain lengths to represent different stages of opsin biosynthesis (c.f. [24]). These integration intermediates can be generated by the translation of truncated mRNAs that lack a stop codon, thus preventing the release of the polypeptide chain from the ribosome [31]. As a consequence, the polypeptide chains are lodged at the ER translocon and may only be released by the removal of the ribosome, for example by puromycin treatment [31]. During our previous studies, we have found that the interpretation of the cross-linking patterns obtained using such integration intermediates can be complicated by the presence of a heterologous population of nascent chains, most likely as a consequence of ribosome stacking on the truncated mRNA [16, 24]. Therefore, an HA epitope tag was introduced at the C-terminus of each integration intermediate to confirm the authenticity of specific adducts [24]. For each of the opsin TM segments analysed below, two or more locations for single cysteine probes were initially trialled. The probe position generating the clearest adducts with known components of the Sec61 complex was then selected for detailed analysis (data not shown).

### TM4 associates transiently with the translocon during opsin biosynthesis

Building on our previous studies, a single cysteine probe was introduced at position 154 within TM4 and a range of integration intermediates from a chain length of 204 residues, at which stage TM4 is predicted to be capable of engaging the ER translocon, to the full length opsin polypeptide (348 residues plus a 9-residue C-terminal HA tag) were generated (Figure 1). Cross-linking was performed with a homobifunctional, sulphhydryl-specific cross-linking reagent BMH, and adducts to subunits of the ER translocon were identified using anti-sera against the Sec61 components, Sec61 $\alpha$  and Sec61 $\beta$ . As previously seen, when longer opsin derived polypeptides were analysed, the population of nascent chains generated was heterogeneous and some species lack an intact C-terminus [24]. Hence we observed differential recognition of translation products by antisera specific to the N- and C-termini of the polypeptide (Fig. 1, c.f. lane 2 with 3, 8 with 9, 14 with 15, 20 with 21 and 26 with 27; see products indicated with asterisks. See also materials and methods section). When the nascent chain is 204 residues long (OP204[cys154]), a discrete cross-linking adduct with Sec61 $\beta$  was observed (Fig. 1, lane 6,  $\beta$ ). This adduct was formed between Sec61 $\beta$  and a genuine OP204 integration intermediate as confirmed by immunoprecipitation with both  $\alpha$ -Op and  $\alpha$ -HA antisera (Fig. 1, c.f. lane 6 to lanes 2 and 3). No authentic adduct with Sec61 $\alpha$  that could be recognized by the  $\alpha$ -HA antibody was seen (Fig. 1, lane 5), and the only product identified is radiolabelled Sec61 $\alpha$  resulted from the translation of endogenous mRNA (Fig. 1, lane 5, closed circle; data not shown), as previously described [24]. When the nascent chain is lengthened to 259 residues, authentic adducts with Sec61 $\beta$  are lost (Fig. 1, lane 12) and further extension of the chain did not result in any authentic adduct formation with Sec61 $\alpha$  or Sec61 $\beta$  (Fig. 1, lanes 17, 18, 23, 24, 29, 30).

When the nascent chain is 259 residues long, TM4 is cross-linked to a ~10 kDa protein that is distinct from Sec61 $\beta$  (Fig. 1, lane 9, P), and this adduct is also observed at longer chain lengths (Fig. 1, lanes 15, 21, 27, products marked P and P?). Although significantly weaker



than previously described, these adducts bear a clear resemblance to the previously described ~10 kDa protein, PAT-10 [16, 24]. Taken together, these data suggest that TM4 enters the Sec61 complex when the nascent chain is 204 residues long but has fully exited the translocon and is most likely in the vicinity of the PAT-10 component at a chain length of 259 residues and beyond. Significantly, a 259 residue opsin chain is predicted to represent a stage of biosynthesis at which TM5 would engage the Sec61 translocon (c.f. Fig. 1, OP259 cartoon).

### TM5 is adjacent to PAT-10 throughout its integration

In order to analyse the integration of the more C-terminal TM domains of opsin, long integration intermediates have to be generated, and such intermediates often form cross-linking adducts that are weak and diffuse [16, 24], with significantly more incomplete chains lacking the C-terminal HA tag (see Fig. 1, cf. lane 26 and 27, products denoted with asterisks). Hence, the results from the cross-linking analysis of such long intermediates can be ambiguous. In order to verify the analysis of cross-linking products obtained using probes located in opsin TM5 to TM7, a shorter version of the opsin polypeptide chain was generated. We noted that residues 195 to 348 of opsin can be expressed and folded independently of the N-terminal portion [32]. Hence, the first 35 residues of opsin containing the two N-glycosylation sites were fused to this C-terminal portion of opsin, comprising TM5 to TM7, to generate the artificial OPN/5-7 polypeptide chain (Fig. 2A). This allowed shorter integration intermediates to be generated for the analysis of TM5, TM6 and TM7 insertion.

The topology of the OPN/5-7 nascent chains (Fig. 2B) was first ascertained by ensuring that all cross-linking adducts were formed with OPN/5-7 integration intermediates which were doubly N-glycosylated. A cysteine probe was introduced at residue 70 (equivalent to residue 229 in the full-length opsin polypeptide chain) at the C-terminal boundary of TM5, and BMH-mediated cross-linking performed for the shortest OPN/5-7 integration intermediate analysed, OPN/5-7 100[cys70]. The resulting products were then subjected to immunoprecipitation and endoglycosidase H treatment. The bulk of the OPN/5-7 100[cys70] chains were endoglycosidase H sensitive (Fig. 2C, c.f. lanes 1 and 6), and formed adducts with Sec61 $\alpha$  and a ~21 kDa putative ribosomal protein [25] that were also both endoglycosidase H sensitive (Fig. 2C, c.f. lanes 2-4 and lanes 7-9,  $\alpha$  and X). This behaviour confirmed that the OPN/5-7 integration intermediates were correctly glycosylated and authentically oriented at the ER translocon (Fig. 2C).

In the context of the OPN/5-7 chains, TM5 is now the first TM domain, TM6 the second and TM7 the third. In order to ensure that the cross-linking results observed were representative of the molecular environment of TM5, TM6 and TM7 in their original position within full-length opsin, a cross-linking analysis with authentic opsin integration intermediates was performed for comparison. Hence, opsin integration intermediates of increasing chain lengths were subjected to a cross-linking analysis and adducts with translocon components were examined. OP259[cys229] gave adducts to both Sec61 $\alpha$  and Sec61 $\beta$  (Fig. 3A, lanes 2-6,  $\alpha$ ,  $\beta$ ), suggesting that TM5 has engaged the translocon at this stage. A currently unidentified adduct of ~21 kDa is also apparent on cross-linking with OP259[cys229] (Fig. 3A lanes 2 and 3, X), which may represent a ribosomal subunit [16]. When longer nascent chains were examined (OP304 to OP357), adducts to translocon components could be observed but as the chain length increases it becomes difficult to be confident that these products are immunoprecipitated with the  $\alpha$ -HA antiserum and therefore represent adducts with authentic integration intermediate (Fig. 3A, lanes 20-24,  $\alpha$ ?,  $\beta$ ?). In contrast to the adducts with the Sec61 subunits, very distinct adducts with a ~10 kDa component were seen with integration intermediates of 304 residues in length and beyond (Fig. 3A, lanes 9, 15 and 21). These

adducts bear a striking resemblance to products previously observed when cross-linking from a probe in TM1 [16], and we conclude that TM5 most likely has a strong association with a PAT-10-like protein during its integration.

When cross-linking analysis was performed for the equivalent OPN/5-7 nascent chains, authentic adducts obtained with translocon components could be more readily distinguished (Fig. 3B). Adducts to Sec61 $\alpha$  and, at chain lengths of 145 residues and beyond, PAT-10 were observed and these adducts were clearly recognized by the  $\alpha$ -HA antibody (Fig. 3B, lanes 2-4, 7-9, 12-14 and 17-19,  $\alpha$ , P). Whilst OPN/5-7 100[cys70] gave a clear adduct with Sec61 $\beta$  (Fig. 3B, lane 5), at longer chain lengths it was difficult to authenticate the Sec61 $\beta$  adduct by immunoprecipitation with the  $\alpha$ -HA antibody due the presence of the much stronger PAT-10 product with a similar molecular weight (Fig. 3B, c.f. lanes 10, 15 and 20,  $\beta$ ). Since adduct formation with Sec61 $\alpha$  persisted until the nascent chain reaches the equivalent of full length (OP198), this indicates that TM5 most likely remains associated with the ER translocon until protein synthesis is complete.

The extended association of opsin TM5 with the Sec61 complex may reflect its specific recognition and retention by the ER translocon. Alternatively, TM5 may remain adjacent to the translocon because the polypeptide tether to the ribosome is simply too short for its complete release from the Sec61 complex to occur. In order to distinguish between these two possibilities, the C-terminus of the OPN/5-7 polypeptide chain was artificially extended with a hydrophilic tail taken from residues 31-220 of bovine preprolactin, generating OPN/5-7-PPL 229[cys70] and OPN/5-7-PPL 280[cys70]. When these polypeptides were analysed by cross-linking, authentic adducts with both Sec61 $\alpha$  and PAT-10 were observed (Fig. 3B, lanes 22-25 and 27-30). Taken as a whole, these results suggest that TM5 remains associated with the ER translocon throughout nascent chain synthesis and that this association persists even when the chain is artificially extended. In addition, TM5 appears to encounter PAT-10 after it moves away from Sec61 $\alpha$ , and it remains associated with PAT-10 throughout opsin nascent chain synthesis.

### The opsin nascent chain is adjacent to a single PAT-10 molecule during biosynthesis

A distinctive feature of TM1 integration is its association with a novel ~10 kDa protein, PAT-10, that is observed as soon as TM2 enters the ER translocon and that remains as long as the polypeptide is ribosome associated [16, 24]. Whilst no cross-linking to PAT-10 from TM2 or TM3 has been observed [24], we find that TM5, and to a much lesser extent TM4, exhibit discrete adduct formation extremely consistent with the partner being PAT-10. Hence, as previously observed for the association of TM1 with PAT-10 [24], these adducts are observed only after the subsequent TM segment has been synthesized (see Figs. 1 and 3A) and they persist until the end of polypeptide synthesis. In order to determine whether the opsin protein interacts with multiple copies of PAT-10, two cysteine probes, one in TM1 (cys56) and the other flanking TM5 (cys229), were introduced into an OP304 integration intermediate. OP304 being the shortest integration intermediate that would permit both TM1 and TM5 to form cross-linking adducts with PAT-10 (see Fig. 3A, lane 9 and Fig. 4, lane 3, P). This approach has been successfully used previously to insert cysteine probes into both TM1 and TM3, to allow simultaneous cross-linking to PAT-10 and Sec61 $\beta$  respectively [24].

Cross-linking analysis of OP304[cys56,229] generated a single major product which corresponds to the polypeptide chain being cross-linked to only one copy of PAT-10 (Fig. 4, lane 6, P). No obvious higher molecular weight adducts were observed, as would be predicted

if TM1 and TM5 were each cross-linked to a different ~10 kDa components. We therefore conclude that TMs 1 and 5 are most likely adjacent to the same PAT-10 molecule during opsin biosynthesis.

### TM6 remains associated with the translocon throughout opsin biosynthesis

The molecular environment of opsin TM6 was examined using nascent opsin chains with a single cysteine residue at position 275 near the C-terminal end of the TM segment. Integration intermediates of chain lengths 304, 339 and 357 residues of the opsin polypeptide chain were generated and examined by site-specific cross-linking (Fig. 5A). The shortest integration intermediate, OP304[cys275], gave a discrete adduct with Sec61 $\beta$ , indicating that TM6 has engaged the translocon at this stage (Fig. 5A, lane 2-6,  $\beta$ ). When the nascent chain reaches 339 residues, TM6 formed adducts with both Sec61 $\alpha$  and Sec61 $\beta$  (Fig. 5A, lanes 8-12,  $\alpha$ ,  $\beta$ ) and these interactions appeared to be maintained until the nascent chain reaches full length at 357 residues, although as previously discussed, the authenticity of adducts formed with longer opsin chain lengths was less definitive (Fig. 5A, lanes 14-18,  $\alpha?$ ,  $\beta?$ ).

In order to authenticate these adducts, a parallel cross-linking analysis was performed with the OPN/5-7 polypeptide chain (Fig. 5B). When the equivalent integration intermediates were analysed, a similar pattern of adduct formation was observed. OPN/5-7 145[cys116] gave a distinct adduct to Sec61 $\beta$  (Fig. 5B, lanes 2-5,  $\beta$ ), while OPN/5-7 180[cys116] and OPN/5-7 198[cys116] produced discrete adducts with both Sec61 $\alpha$  and Sec61 $\beta$  (Fig. 5B, lanes 7-10 and 12-15,  $\alpha$ ,  $\beta$ ). These results indicate that opsin TM6 in the OPN/5-7 polypeptide chain has engaged the Sec61 complex when the nascent chain is 145 residues long and it remains associated with the translocon throughout synthesis.

In order to further examine the nature of the association between opsin TM6 and the Sec61 complex, the polypeptide chain tether was further extended to generate OPN/5-7-PPL 229[cys116] and OPN/5-7-PPL 280[cys116]. Cross-linking of these nascent chains generated adducts with Sec61 $\alpha$  (Fig. 5B, lanes 17-20 and 22-25,  $\alpha$ ) consistent with opsin TM6 remaining in close proximity to the Sec61 complex even when TM7 has been fully synthesised and the C-terminus of the polypeptide artificially extended.

### TM7 engages the translocon throughout opsin biosynthesis

In order to examine the integration of the final opsin TM segment, TM7, cross-linking was performed using integration intermediates with a cysteine probe at position 287 near the N-terminus of TM7. Cross-linking from both the OP339[cys287] and OP357[cys287] integration intermediates produced adducts with Sec61 $\beta$  and putative adducts with Sec61 $\alpha$  (Fig. 6A, lanes 2-5 and 7-10,  $\beta$ ,  $\alpha?$ ). To further examine opsin TM7 insertion, and to confirm the authenticity of adducts between TM7 and Sec61 $\alpha$ , equivalent integration intermediates of the OPN/5-7 polypeptide chain were also analysed by cross-linking. Adducts with Sec61 $\beta$  were observed for both OPN/5-7 180[cys128] and OPN/5-7 198[cys128] as was seen with equivalent opsin integration intermediates in previous experiments (Fig. 6B, lanes 2-5 and 7-10,  $\beta$  c.f. with Fig. 6A, lanes 2-5 and 7-10,  $\beta$ ). In addition, adducts with Sec61 $\alpha$  were clearly identified for these nascent chains (Fig. 6B, lanes 2-4 and 7-9,  $\alpha$ ). As with the analysis of opsin TM5 and TM6, two additional integration intermediates with extended chain lengths were examined in order to determine if TM7 could exit the translocon completely when the polypeptide tether is lengthened. The integration intermediates OPN/5-7-PPL 229[cys128] and OPN/5-7-PPL 280[cys128] both produced adducts with Sec61 $\alpha$  and Sec61 $\beta$  (Fig. 6B,

lanes 12-15, 17-19,  $\alpha$ ,  $\beta$ ). The adducts with Sec61 $\alpha$  and Sec61 $\beta$  observed at the longest length of 280 residues (Fig. 6B, lanes 17-20,  $\alpha$ ,  $\beta$ ) were not present when integration intermediates were treated with puromycin to release the opsin polypeptide chains from the ribosome and enable complete integration and exit from the Sec61 translocon prior to cross-linking (Fig. 6B, lanes 22-25). Opsin TM7 therefore remains associated with the Sec61 translocon well beyond the maximum chain length for wild type opsin and is only fully integrated into the ER membrane after artificial termination of translation

### Retention at the Sec61 translocon is TM-specific

The prolonged retention of opsin TM7 at the ER translocon may either reflect the intrinsic properties of the TM segment itself or may alternatively be a consequence of its relative position as the final TM segment of opsin to encounter the translocon. In order to try and distinguish between these two possibilities, we initially attempted to analyse the membrane integration of opsin TM7 alone without any other opsin TMs being present. However, we found that opsin TM7 is an extremely poor ER targeting signal as has been previously suggested [33] and were unable to pursue this approach (data not shown).

We therefore devised an alternative approach to establish the basis for the retention of TM7 at the Sec61 translocon by replacing it with an alternative TM domain within the context of the OPN/5-7-PPL construct. On the basis of its rapid exit from the ER translocon [24], we identified opsin TM3 as a good candidate to replace TM7. Our previous study had analysed the authentic opsin polypeptide and we first established whether TM3 was able to efficiently exit the Sec61 translocon in isolation. To this end, OPN/TM3PPL was generated comprising the opsin N-terminal region with two sites for N-glycosylation, TM3 of opsin, the C-terminal cytoplasmic tail of opsin, and finally a large segment of PPL to allow artificially long integration intermediates to be generated (Supplementary Fig. 1A). When analysed in vitro, OPN/TM3PPL is efficiently glycosylated at all chain lengths tested (Supplementary Fig. 1B, lanes 1-3, 6-8, 11-13 and 16-18, open arrowhead) and on the basis of a careful cross-linking analysis we conclude that TM3 is efficiently partitioned into the lipid phase of the ER membrane (Supplementary Fig. 1B), as indicated by our previous studies on early events in the biosynthesis and integration of opsin [24].

Having demonstrated the intrinsic capacity of opsin TM3 to efficiently exit the Sec61 translocon, both in its 'native' location between TM2 and TM4 [24] and also in isolation from other opsin TM regions (Supplementary Fig. 1), we now wished to test the behaviour of TM3 when it was used to replace TM7 in the OPN/5-7-PPL polypeptide. The OPN/TM5,6,3-PPL construct was therefore generated (Fig. 7A), and cross-linking was performed from a single cysteine at residue 135, equivalent to residue 115 in full-length opsin [24] and residue 44 in OPN/TM3PPL (see supplementary Fig. 1). Cross-linking from the OPN/TM5,6,3-PPL 192[cys135] integration intermediate produced clear adducts with Sec61 $\alpha$  and Sec61 $\beta$  (Fig. 7B, lanes 2-5,  $\alpha$  and  $\beta$ ) indicating that it has fully engaged with the ER translocon. Extension of the integration intermediates to 210 residues resulted in very weak adduct formation (Fig. 7B, lanes 7-10,  $\alpha$ ?,  $\beta$ ?), and by 241 and 292 residues no adducts with Sec61 subunits were detected (Fig. 7B, lanes 12-15 and data not shown). On the basis of the length of the C-terminal tether after the end of the final TM domain, OPN/TM5,6,3PPL 210 is equivalent to OPN/5-7-PPL 198 and OPN/TM5,6,3PPL 241 is equivalent to OPN/5-7-PPL 229. When the cross-linking profiles of these intermediates are compared it is clear that TM3 has exited the translocon at a stage where TM7 has not (c.f. Fig. 6B, lanes 11-15 with Fig. 7B, lanes 11-15). We therefore conclude that the prolonged association of TM7 with the Sec61 translocon is not due to its location as the final TM domain within the polypeptide chain.

## DISCUSSION

We have used site-specific chemical cross-linking to characterise the molecular environment of TM4 to TM7 of opsin during its biosynthesis in order to study how different TM regions are accommodated by the Sec61 translocon during ER integration. By using immunoprecipitation from both the N- and C-termini of the nascent polypeptide chains, we can identify adducts with defined integration intermediates [16, 24]. Thiol specific, homobifunctional cross-linking reagents such as BMH, have been successfully applied to the study of protein biosynthesis at the ER [25] and the mitochondrion [34]. When combined with the use of engineered single cysteine residues these reagents provide a platform for a site-specific analysis of the environment of individual TM segments at defined stages of biosynthesis [16, 24]. However, this technique can only identify components adjacent to the nascent polypeptide that contain an accessible cysteine. This means that, in the context of the data presented here, any loss of adduct formation observed upon increasing the length of the nascent opsin chain may in principal reflect either a lateral exit of the TM segment into the lipid bilayer and away from the ER translocon site, or a conformational change that renders one or more cysteine residues inaccessible. In practice, a loss of cross-linking to the Sec61 complex is considered to be a reliable report of TM segment integration when either a thiol-mediated approach, or less restricted photocross-linking techniques, are exploited [16, 18, 19, 21, 23, 24, 26, 35].

Unlike the synchronized integration of TM1 to TM3 [24], we find that opsin TM4 engages the Sec61 complex in a relatively straightforward manner as TM3 departs and is itself subsequently 'displaced' as TM5 emerges from the ribosome. In contrast, TM5 remains in close proximity to the ER translocon until the entire opsin polypeptide chain is synthesised. A further level of complexity with regard to the analysis of TM5 is that long integration intermediates have to be analysed, and such chains generally give low efficiency adducts that can be difficult to unambiguously confirm as being authentic. To validate such cross-linking data, a shortened construct containing only the N-terminal region and TM5 to TM7 of opsin (OPN/5-7) was generated. Since the relative position TM segments was altered i.e., TM5 becomes the first TM segment etc., the resulting data should be interpreted carefully. Hence, wherever feasible, any cross-linking data obtained using OPN/5-7 was compared with that from equivalent authentic opsin intermediates. This approach allowed an unambiguous assignment of the cross-linking data obtained.

When analysed in this way, a striking feature of TM5 is its constant association with the Sec61 complex throughout nascent chain synthesis. This association was apparent with both authentic opsin and OPN/5-7, and suggests that, unlike TMs 1 to 4 [16, 24, see also this study], TM5 remains associated with the ER translocon until the complete opsin chain is synthesised. This behaviour is clearly significant and implies that the Sec61 complex can distinguish between different TM segments and may regulate their release into the ER membrane. The distinct association of different opsin TM segments with the ER translocon also supports the idea that TM regions display sequence-dependent interactions with the Sec61 complex [19].

As for TM5, the environment of TM6 and TM7 was examined using both authentic opsin chains, and the shorter OPN/5-7 derivatives. In both of these contexts, TM6 was found associated with the Sec61 complex from the shortest integration intermediate examined to the full-length polypeptide chain, indicating that TM6 engages the translocon for a prolonged period of time. Like TM6, the cross-linking pattern of TM7 in the context of both authentic opsin and OPN/5-7 derived chains showed that TM7 also remains at the translocon

throughout nascent chain synthesis. Even when integration intermediates were artificially extended by the addition of preprolactin polypeptide at the C-terminus (OPN/5-7-PPL), TMs 5, 6 and 7 remained associated with the translocon. The loss of cross-linking between opsin TMs 4 to 7 and the Sec61 $\beta$  subunit detected upon chain extension are entirely consistent with the relocation of these TMs within the Sec61 complex during biosynthesis as previously suggested [16, 24, see Fig. 8].

Another striking feature of opsin TM5 is its strong adduct formation with PAT-10, akin to that previously observed with TM1 [c.f. 24]. Whilst it is a formal possibility that TM5 is cross-linked to a different ~10 kDa component, we consider this unlikely, particularly since only one adduct with a ~10 kDa protein was observed when cross-linking was performed with opsin chains containing two cysteine residues capable of adduct formation with PAT-10, one in TM1 and the other in TM5. Furthermore, TM5 forms clear adducts with PAT-10 when we analyse authentic opsin nascent chains, ruling out an artificial association with OPN/5-7. We previously suggested that PAT-10 may be an accessory factor or 'TM chaperone' that functions to regulate the release of specific TM segments from the ER translocon into the membrane [16]. Intriguingly, in our current study we observe a prolonged association of PAT-10 with TM1 and TM5, each of which represent the first TM segment of two well characterised opsin sub-domains capable of independent folding and assembly [32, 36]. Hence, we propose that PAT-10 may coordinate the exit of biosynthetically distinct sub-domains of polytopic membrane proteins such as opsin (Fig. 8). As for TRAM [19, 37], the location of PAT-10 at the translocon periphery [16, 24] supports a model where TM segment integration into the ER membrane is a highly regulated process (see Fig. 8).

The extended interaction between opsin TMs 5 to 7 and the Sec61 complex led us to investigate the basis for TM7 retention. As previously described [33], we were unable to analyse TM7 alone (data not shown), most likely because it is the shortest and least hydrophobic transmembrane helix in opsin [38, 39]. In contrast, when TM3 was studied in isolation, using OPN/TM3PPL (Supplementary Fig. 1A), the resulting integration intermediates were efficiently glycosylated at all chain lengths tested, and cross-linking showed TM3 rapidly exited the ER translocon upon chain extension. Most significantly, when TM7 was replaced with TM3 to generate OPN/TM5,6,3-PPL, the behaviour of the final TM region was altered substantially, with TM3 clearly exiting the translocon at a much earlier stage of biosynthesis. Hence, it is the properties of the TM segment and not its relative location within the polypeptide chain that result in the prolonged association of opsin TM7 with the ER translocon (Fig. 8).

The clear distinction in the behaviour of opsin TM3 and TM7 at the ER translocon correlates well with previous biophysical studies that demonstrate the importance of amino acid hydrophobicity during the partitioning of TM helices into the lipid phase of the ER membrane [20]. On the basis of this study we conclude that it is the Sec61 complex, together with accessory factors, that directly determines the individual fate of specific TM segments. The retention of the weakly hydrophobic TM7 in the translocon compared to the rapid lateral exit of TM3 into the ER membrane supports a model in which nascent TM segments entering the ER translocon may be rapidly released or selectively retained depending upon their specific properties and assembly requirements. Given the function of the Sec61 complex as a molecular machine that introduces specific polypeptide regions into a phospholipid bilayer one can rationalise the need for it to be equipped to deal with non-optimal TM segments. In the case of opsin TM7, hydrophobicity must be significantly compromised by the presence of lys296, which is critical for chromophore attachment and hence biological function [38, 39].

## TM segment delay at the ER translocon

Hence, the release of opsin TM7 from the translocon is most likely contingent upon the remainder of the polypeptide assuming a suitable conformation to accommodate it (Fig. 8).

In summary, our detailed analysis of opsin biogenesis supports a model where the ER translocon specifically regulates the integration of the TM segments within a polytopic membrane protein so as to accommodate the folding and functional requirements of the substrate itself. Hence, the selective retention of multiple TM domains at distinct sites within the translocon may be required in order to allow the correct assembly of complex polytopic proteins [24, 26]. The further consolidation of this model will require a detailed analysis of several different precursors, and a greater structural understanding of the active ER translocon.

## ACKNOWLEDGMENTS

We thank Martin Pool and Blanche Schwappach for helpful input during the preparation of this manuscript. This work was supported by a Wellcome Trust PhD award (N.I.) and grant (S.G.C.) and PhD (B.C.S.C.) funding from the BBSRC. We thank Bernhard Dobberstein, Paul Hargrave and Richard Zimmermann for providing antisera used during this study.

Stage 2(a) POST-PRINT



## REFERENCES

- 1 Alder, N. N. and Johnson, A. E. (2004) Cotranslational membrane protein biogenesis at the endoplasmic reticulum. *J. Biol. Chem.* **279**, 22787-22790
- 2 Lecomte, F. J., Ismail, N. and High, S. (2003) Making membrane proteins at the mammalian endoplasmic reticulum. *Biochem. Soc. Trans.* **31**, 1248-1252
- 3 Ellgaard, L. and Helenius, A. (2003) Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* **4**, 181-191
- 4 Osborne, A. R., Rapoport, T. A. and van den Berg, B. (2005) Protein Translocation By The Sec61/Secy Channel. *Annu. Rev. Cell. Dev. Biol.* **21**, 529-550
- 5 Van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. and Rapoport, T. A. (2004) X-ray structure of a protein-conducting channel. *Nature*. **427**, 36-44
- 6 Junne, T., Schwede, T., Goder, V. and Spiess, M. (2006) The plug domain of yeast Sec61p is important for efficient protein translocation, but is not essential for cell viability. *Mol. Biol. Cell.* **17**, 4063-4068
- 7 Maillard, A. P., Lalani, S., Silva, F., Belin, D. and Duong, F. (2007) Deregulation of the SecYEG translocation channel upon removal of the plug domain. *J. Biol. Chem.* **282**, 1281-1287
- 8 Li, W., Schulman, S., Boyd, D., Erlandson, K., Beckwith, J. and Rapoport, T. A. (2007) The plug domain of the SecY protein stabilizes the closed state of the translocation channel and maintains a membrane seal. *Mol. Cell.* **26**, 511-521
- 9 Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C. L., 3rd, Ban, N. and Frank, J. (2005) Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature*. **438**, 318-324
- 10 Osborne, A. R. and Rapoport, T. A. (2007) Protein translocation is mediated by oligomers of the SecY complex with one SecY copy forming the channel. *Cell.* **129**, 97-110
- 11 Beckmann, R., Spahn, C. M., Eswar, N., Helmers, J., Penczek, P. A., Sali, A., Frank, J. and Blobel, G. (2001) Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell.* **107**, 361-372
- 12 Menetret, J. F., Hegde, R. S., Heinrich, S. U., Chandramouli, P., Ludtke, S. J., Rapoport, T. A. and Akey, C. W. (2005) Architecture of the ribosome-channel complex derived from native membranes. *J. Mol. Biol.* **348**, 445-457
- 13 Wirth, A., Jung, M., Bies, C., Frie, M., Tyedmers, J., Zimmermann, R. and Wagner, R. (2003) The Sec61p complex is a dynamic precursor activated channel. *Mol. Cell.* **12**, 261-268
- 14 Do, H., Falcone, D., Lin, J., Andrews, D. W. and Johnson, A. E. (1996) The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell.* **85**, 369-378
- 15 Heinrich, S. U., Mothes, W., Brunner, J. and Rapoport, T. A. (2000) The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell.* **102**, 233-244
- 16 Meacock, S. L., Lecomte, F. J., Crawshaw, S. G. and High, S. (2002) Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. *Mol. Biol. Cell.* **13**, 4114-4129
- 17 von Heijne, G. (2006) Membrane-protein topology. *Nat. Struct. Mol. Biol.* **7**, 909-918
- 18 Martoglio, B., Hofmann, M. W., Brunner, J. and Dobberstein, B. (1995) The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. *Cell.* **81**, 207-214
- 19 McCormick, P. J., Miao, Y., Shao, Y., Lin, J. and Johnson, A. E. (2003) Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. *Mol. Cell.* **12**, 329-341
- 20 Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H. and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*. **433**, 377-381
- 21 Heinrich, S. U. and Rapoport, T. A. (2003) Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. *EMBO J.* **22**, 3654-3663
- 22 Buck, T. M., Wagner, J., Grund, S. and Skach, W. R. (2007) A novel tripartite motif involved in aquaporin topogenesis, monomer folding and tetramerization. *Nat. Struct. Mol. Biol.* **14**, 762-769
- 23 Sauri, A., Saksena, S., Salgado, J., Johnson, A. E. and Mingarro, I. (2005) Double-spanning plant viral movement protein integration into the endoplasmic reticulum membrane is signal recognition particle-dependent, translocon-mediated, and concerted. *J. Biol. Chem.* **280**, 25907-25912
- 24 Ismail, N., Crawshaw, S. G. and High, S. (2006) Active and passive displacement of transmembrane domains both occur during opsin biogenesis at the Sec61 translocon. *J. Cell Sci.* **119**, 2826-2836

## TM segment delay at the ER translocon

- 25 Laird, V. and High, S. (1997) Discrete cross-linking products identified during membrane protein biosynthesis. *J. Biol. Chem.* **272**, 1983-1989
- 26 Sadlish, H., Pintonzo, D., Johnson, A. E. and Skach, W. R. (2005) Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nat. Struct. Mol. Biol.* **12**, 870-878
- 27 Dale, H., Angevine, C. M. and Krebs, M. P. (2000) Ordered membrane insertion of an archaeal opsin in vivo. *Proc. Natl. Acad. Sci. USA.* **97**, 7847-7852
- 28 Cui, L., Aleksandrov, L., Chang, X. B., Hou, Y. X., He, L., Hegedus, T., Gentsch, M., Aleksandrov, A., Balch, W. E. and Riordan, J. R. (2007) Domain interdependence in the biosynthetic assembly of CFTR. *J. Mol. Biol.* **365**, 981-994
- 29 Adamus, G., Zam, Z. S., Arendt, A., Palczewski, K., McDowell, J. H. and Hargrave, P. A. (1991) Anti-rhodopsin monoclonal antibodies of defined specificity: characterization and application. *Vision Res.* **31**, 17-31
- 30 Wilson, R., Allen, A. J., Oliver, J., Brookman, J. L., High, S. and Bulleid, N. J. (1995) The translocation, folding, assembly and redox-dependent degradation of secretory and membrane proteins in semi-permeabilized mammalian cells. *Biochem. J.* **307**, 679-687
- 31 Gilmore, R., Collins, P., Johnson, J., Kellaris, K. and Rapiejko, P. (1991) Transcription of full-length and truncated mRNA transcripts to study protein translocation across the endoplasmic reticulum. In *Methods in Cell Biology* (Tartakoff, A. M., ed.) pp. 223-239., Academic Press
- 32 Ridge, K. D., Lee, S. J. S. and Abdulaev, N. G. (1996) Examining rhodopsin folding and assembly through expression of polypeptide fragments. *J. Biol. Chem.* **271**, 7860-7867
- 33 Audigier, Y., Friedlander, M. and Blobel, G. (1987) Multiple topogenic sequences in bovine opsin. *Proc. Natl. Acad. Sci. USA.* **84**, 5783-5787
- 34 Endres, M., Neupert, W. and Brunner, M. (1999) Transport of the ADP/ATP carrier of mitochondria from the TOM complex to the TIM22.54 complex. *Embo J.* **18**, 3214-3221
- 35 Mothes, W., Heinrich, S. U., Graf, R., Nilsson, I., von Heijne, G., Brunner, J. and Rapoport, T. A. (1997) Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell.* **89**, 523-533
- 36 Ridge, K. D., Lee, S. J. S. and Yao, L. L. (1995) In vivo assembly of rhodopsin from expressed polypeptide fragments. *Proc. Natl. Acad. Sci. USA.* **92**, 3204-3208
- 37 Voigt, S., Jungnickel, B., Hartmann, E. and Rapoport, T. A. (1996) Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J. Cell Biol.* **134**, 25-35
- 38 Palczewski, K. (2006) G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **75**, 743-767
- 39 Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M. and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science.* **289**, 739-745

## FIGURE LEGENDS

### Fig. 1. TM4 exits the Sec61 translocon upon chain extension.

A schematic of ribosome-bound integration intermediates is presented above each set of products. The dashed line represents the transmembrane (TM) segment, the solid line indicates the hydrophilic portion of the chain and the star denotes the cysteine probe within TM4. Each integration intermediate was synthesized in a rabbit reticulocyte lysate translation system in the presence of [<sup>35</sup>S]met/cys and digitonin-permeabilised mammalian HT-1080 cells. Membrane-associated integration intermediates were isolated by centrifugation and cross-linking performed by the addition of the homobifunctional cross-linking reagent, BMH (+), or mock-treated with solvent only (-). Opsin-derived nascent chains were immunoprecipitated using either a monoclonal antibody specific for the N-terminus of the polypeptide (Op), for the HA epitope on the C-terminus of the integration intermediates (HA), or with a non-related serum (NS). For each chain length, authentic doubly N-glycosylated opsin chains are indicated by an open arrowhead and their non-glycosylated counterparts by an open circle. Shorter opsin chains that lack the C-terminal HA epitope tag are indicated by asterisks. Adducts to Sec61 $\alpha$  and Sec61 $\beta$  were identified by immunoprecipitation with appropriate antisera (denoted  $\alpha$  and  $\beta$ ). P indicates adducts to an unidentified protein of ~10 kDa, denoted PAT-10 (lanes 9, 15, 21 and 27). Sec61 $\alpha$  molecules produced by the translation of endogenous mRNA [24] are denoted with closed circles. Vertical black lanes indicate sets of samples resolved on different SDS-polyacrylamide gels. Positions of molecular size markers are indicated to the left.

### Fig. 2. OPN/5-7 opsin deletion mutant is properly inserted into the ER membrane.

(A) A schematic diagram of the OPN/5-7 polypeptide chain. The first 35 amino acids of opsin were fused to a C-terminal portion of opsin (residues 195 to 348), containing TM5 to TM7. The TM segments are shown in black. (B) The predicted topology of the OPN/5-7 polypeptide chain. The TM segments are shown in grey and the numbers indicate the position of TM5 to TM7 with respect to the OPN/5-7 protein chain. N-glycosylated residues in the ER lumen are represented by branched structures. (C) The OPN/5-7 construct contains asparagine-linked glycan groups. An integration intermediate with a length of 100 amino acid residues and a cysteine probe in TM5 (residue 70 of the polypeptide chain) was subjected to cross-linking and immunoprecipitation as described in Fig. 1. Duplicate samples were treated with endoglycosidase H (+ Endo H) to cleave N-linked glycans. Cross-linking adducts with Sec61 $\alpha$  and other unidentified molecules are indicated with ' $\alpha$ ' and 'X' respectively. All other symbols are as defined in the legend to Fig. 1.

### Fig. 3. Opsin TM5 exhibits prolonged association with Sec61 components and PAT-10.

(A) BMH-dependent cross-linking was performed using opsin integration intermediates of 259, 304, 339 and 357 residues all containing a single cysteine probe at position 229 within TM5. (B) OPN/5-7 integration intermediates with lengths of 100, 145, 180 and 198 amino acids, all containing a cysteine at position 70 within TM5 were translated as before. Due to the 159-residue fragment deleted from opsin to generate the OPN/5-7 construct, the four shortest integration intermediate lengths used for the OPN/5-7 construct correlate to the four lengths used in Fig. 3A. Thus, in constructs OP259[cys229] and OPN/5-7 100[cys70], the cysteine is in the same relative position in the TM segment in both constructs, and the TM

## TM segment delay at the ER translocon

segment is the same distance away from the ribosome in both cases. The OPN/5-7 construct was also further extended in length by splicing the soluble region from bovine preprolactin onto the C-terminus of the opsin chain, thus enabling the chain length to be extended up to 229 and 280 residues (lanes 21-30). Putative adducts with Sec61 $\alpha$  and Sec61 $\beta$  are labelled with ' $\alpha$ ' and ' $\beta$ ' respectively. Two unidentified adducts are indicated with 'X' for an adduct potentially with a ribosomal subunit [16] and a 'Y', whilst all other symbols are as previously defined.

**Fig. 4. TM1 and TM5 are adjacent to a single PAT-10 molecule.**

An opsin integration intermediate containing two cysteine probes, the first in TM1 (residue 56) and the second in TM5 (residue 229), was analysed at a previously characterised chain length (OP304) and compared to an integration intermediate of the same length containing only a single cysteine probe in TM1 (OP304[cys56]). All symbols are as previously defined.

**Fig. 5. TM6 remains in close proximity to the translocon throughout opsin biosynthesis.**

(A) Membrane associated integration intermediates of OP304 to OP357 containing a single cysteine probe in TM6 (residue 275) were treated with BMH and the resulting adducts analysed following immunoprecipitation as previously described. (B) The corresponding OPN/5-7 equivalents (OPN5-7 145 to 198[cys116]) were analysed as for panel A. As in Fig. 3, the lengths of these OPN/5-7 integration intermediates were extended by splicing the soluble region of bovine preprolactin onto the C-terminus (OPN/5-7-PPL 229 and 280[cys115], lanes 16-25). Products labelled ' $\alpha\beta$ ' are due to adducts which contain both Sec61 $\alpha$  and Sec61 $\beta$ , whilst all other symbols are as previously defined.

**Fig. 6. TM7 remains associated to the Sec61 complex for a prolonged period.**

(A) Radiolabelled integration intermediates of OP339 and OP357 were cross-linked to adjacent ER components from cysteine residue 287 in TM7 using BMH. (B) The corresponding OPN/5-7 equivalents (OPN5-7 180 and 198[cys128]) were analysed as for panel A. As in Figs. 3 and 5, the lengths of these OPN/5-7 integration intermediates were extended by splicing the soluble region of bovine preprolactin onto the C-terminus (OPN/5-7-PPL 229 and 280[cys128], lanes 11-25). Treatment of OPN/5-7-PPL 280[cys128] with puromycin (+ PURO, lanes 21-25) caused release of the nascent chain from the ribosome, enabling complete integration and exit from the Sec61 translocon as judged by cross-linking (lanes 24 and 25). All symbols are as previously defined.

**Fig. 7. The final TM of opsin integrates more readily with increased hydrophobicity.**

(A) A schematic diagram of the OPN/TM5,6,3PPL polypeptide chain. The first 35 amino acids of opsin were fused to a C-terminal portion of opsin containing TM5, TM6 followed by TM3 and the soluble preprolactin extension. The TM segments are shown in black. (B) BMH-dependent cross-linking was performed using OPN/TM5,6,3PPL integration intermediates of 192, 210 and 241 residues all containing a single cysteine probe at position 135 within TM3 (the final TM segment of the polypeptide chains), and the resulting adducts analysed following immunoprecipitation as previously described. All other symbols are as previously defined.

**Fig. 8. Model for the integration of opsin TMs 4-7.**

A schematic view from the cytosolic face of the ER membrane showing the presumptive 'open' conformation of the Sec61 translocon, based on an archeal complex [5]. Only the  $\alpha$  and  $\beta$  subunits are represented since these are the cross-linking partners of the opsin integration intermediates studied, whilst the PAT-10 protein is as previously described [16, 24]. The relative location of each opsin TM segment (numbered circle) is indicated at different stages of the integration process as deduced from the cross-linking patterns of specific integration intermediates. TM4 is part of an N-terminal independently folding sub-domain of opsin and rapidly exits the Sec61 complex (i-iii). TMs 5, 6 and 7, which can form a C-terminal independently folding sub-domain, remain associated with the translocon (ii-iv) until biosynthesis is complete (v). Data for TMs 1-3 (dashed circles) was generated previously [24]. Both opsin TM1 and TM5 exhibit a pronounced and prolonged association with PAT-10, suggesting a co-ordination step during the biosynthesis of opsin in which two folding sub-domains (TMs 1-4 and TMs 5-7) are integrated independently. PAT-10 may contribute directly to the regulation of this distinct, and substrate specific, membrane integration process.

Figure 1

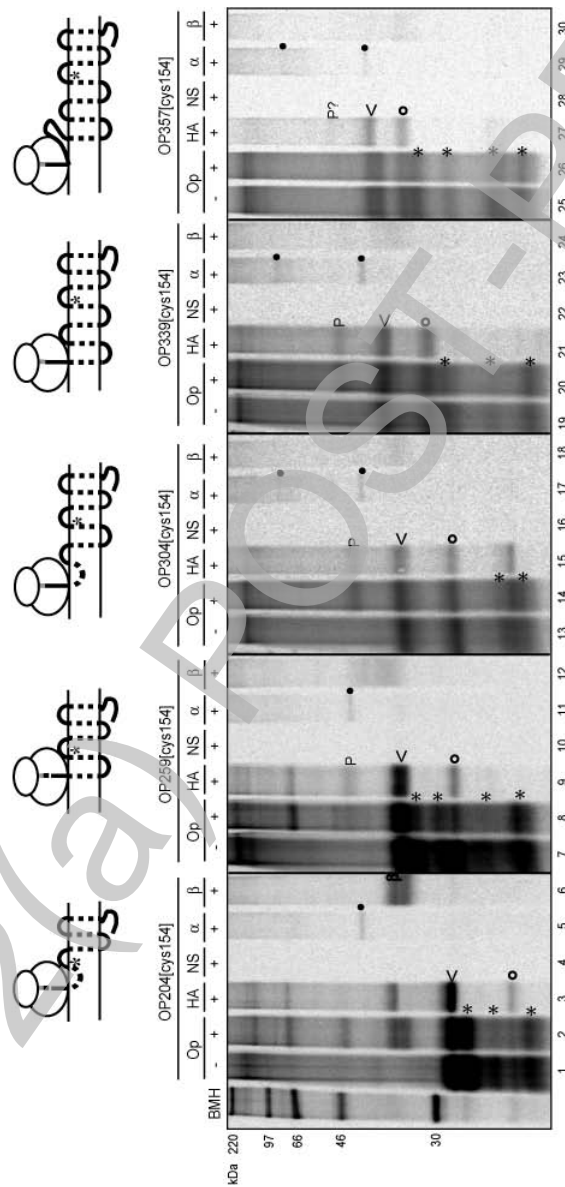


Figure 2

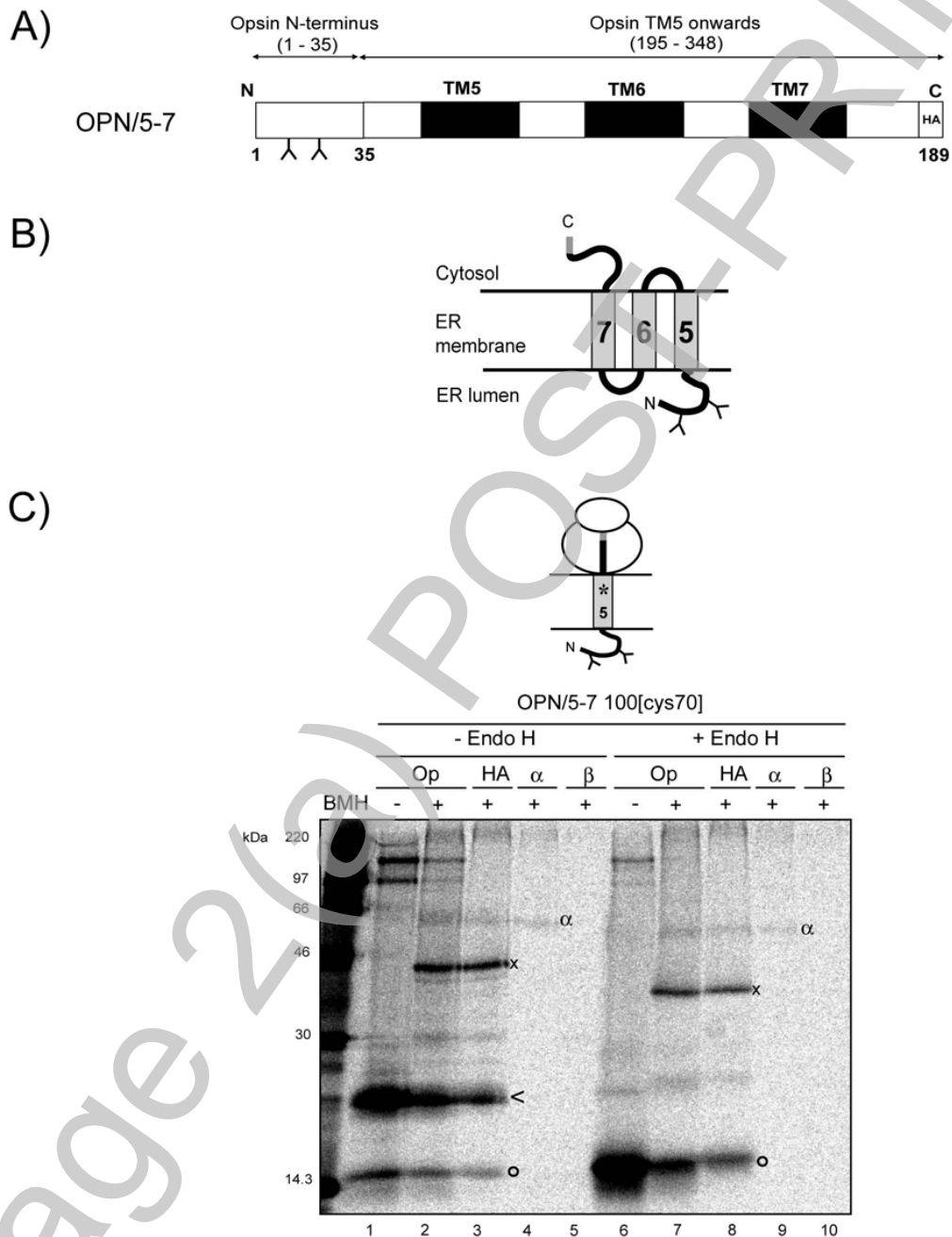






Figure 4

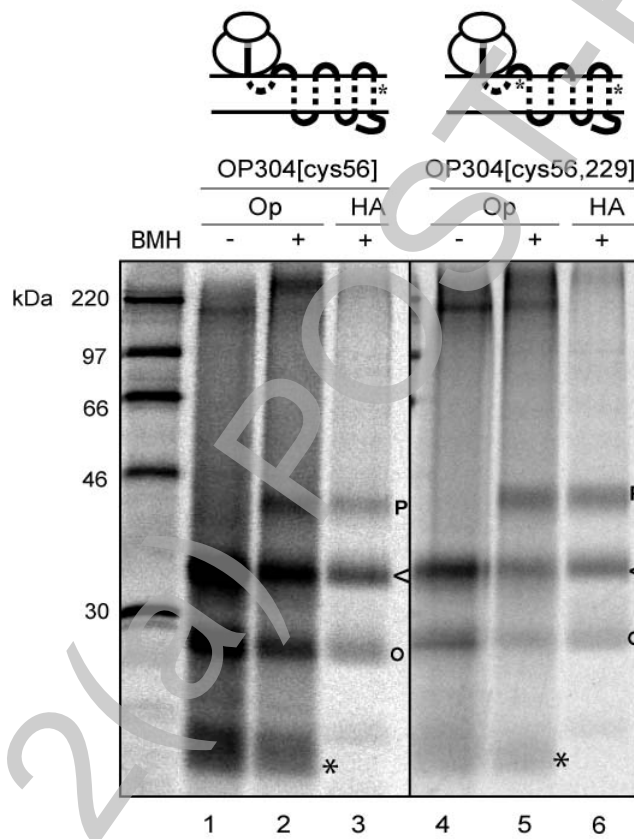


Figure 5

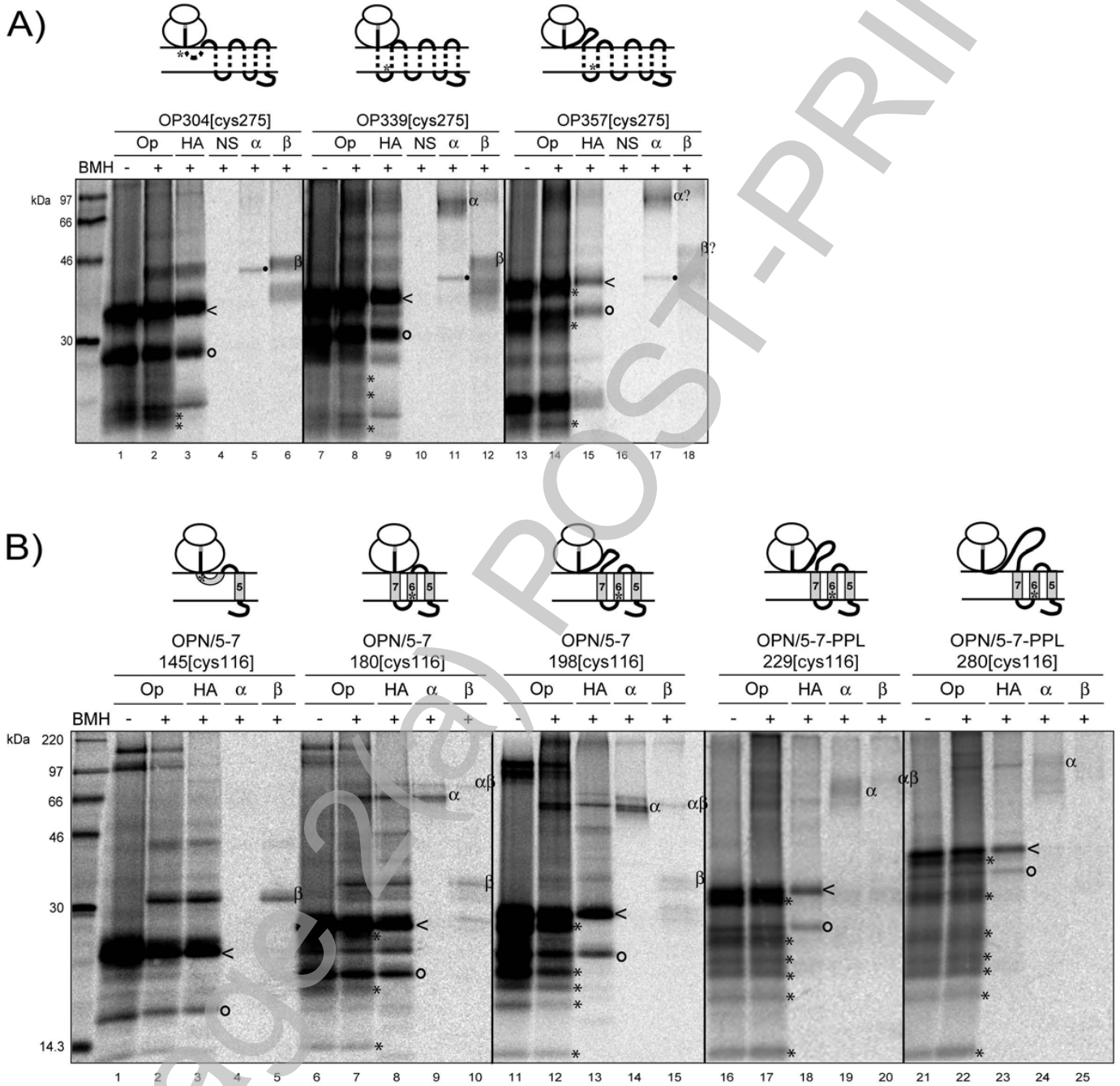


Figure 6

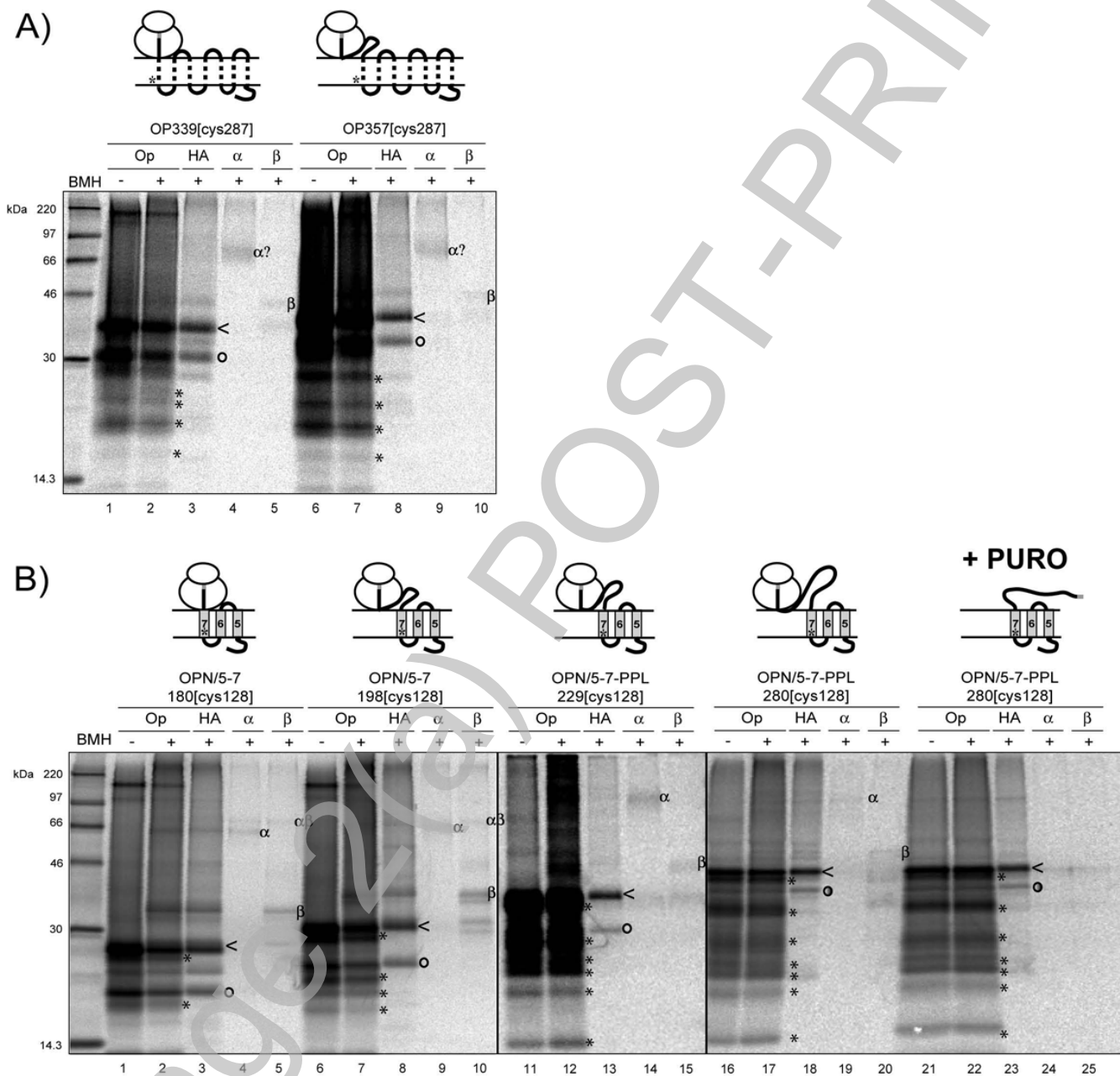
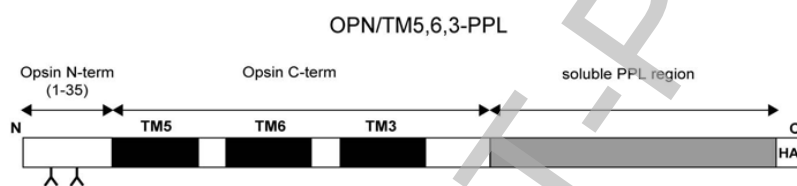


Figure 7

A)



B)

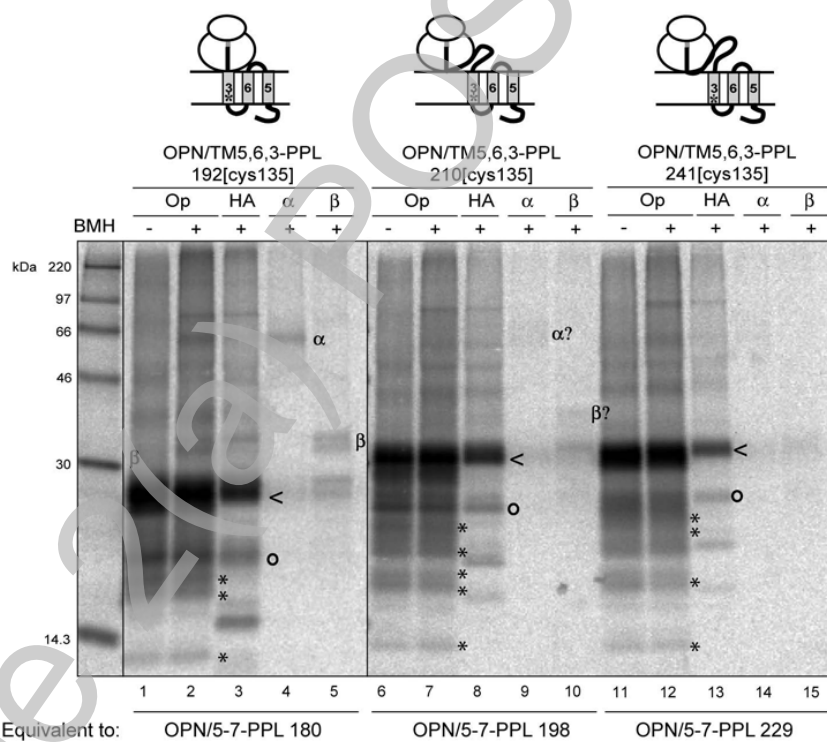


Figure 8

