Topology studies with biosynthetic fragments identify interacting transmembrane regions of the human red-cell anion exchanger (band 3; AE1)

Jonathan D. GROVES¹ and Michael J. A. TANNER

Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, U.K.

The red-cell anion exchanger (band 3; AE1) is a multispanning membrane protein that traverses the bilayer up to 14 times and is N-glycosylated at Asn-642. We have shown that the integrity of six different loops are not essential for stilbene disulphonate-sensitive chloride uptake in *Xenopus* oocytes. We used an N-glycosylation mutagenesis approach to examine the orientation of the N-terminus and the endogenous glycosylation site of each C-terminal fragment by cell-free translation. The fragments initiating in the loops preceding spans 2, 9 and 11 did not insert into the membrane with the expected orientation. Furthermore, N-glycosylation of Asn-642 might facilitate the membrane integration of span 7. The correct integration of spans 2–3 required the presence of the region containing span 4 and that the luminal exposure of the C-terminus of span 7 is increased in the presence

of the region including span 6 or span 8. The results suggest the span 8 region is required for the correct folding of spans 9–10, at least in the presence of the span 11–12 region. Our results suggest that there are intramolecular interactions between the regions of transmembrane spans 1 and 2, 2 and 4, 4 and 5, 7 and 8, 8 and 9–10, and 9–10 and 11–12. Spans 1, 4, 5, 6 and 8 might act as a scaffold for the assembly of spans 2–3, 7 and 9–10. This approach might provide a general method for dissecting the interactions between membrane-spanning regions of polytopic membrane proteins.

Key words: erythrocyte, membrane, membrane protein structure, transmembrane spans, transport.

INTRODUCTION

The human red-cell anion exchanger (band 3; AE1) is a multifunctional polytopic membrane transporter that has been studied extensively as a model for this class of integral proteins (reviewed in [1–3]). It comprises two domains that are structurally and functionally distinct: the 43 kDa N-terminal cytoplasmic domain (residues 1–359) is involved in functions unrelated to anion transport, whereas the 52 kDa C-terminal integral membrane domain (b3mem; residues 360–911) is both necessary and sufficient to mediate the chloride–bicarbonate exchange function [4–6]. The membrane domain is thought to traverse the bilayer 12 or 14 times [2,7–11] and has been shown to contain a high proportion of α -helical structure [12]. Two-dimensional crystallography has yielded a low-resolution structural model of the dimeric membrane domain [13] indicating the presence of three subdomains within the monomer and shows that the cross-sectional area of the molecule in the plane of the membrane is sufficient to permit the packing of up to 14 α -helical transmembrane (TM) spans.

In previous work, we have reported that functionally active band 3 can be expressed in the plasma membrane of *Xenopus* oocytes from pairs of complementary fragments separated within the fourth or sixth cytoplasmic loops and the third or fourth extracellular (EC) loops, but not in the first and second EC loops or with any of the individual fragments [14,15]. In subsequent studies, we showed that several combinations of three or four complementary fragments (i.e. simultaneously divided in two or

three of these permissive loops) co-assemble into the functional anion transporter [16]. These results, combined with evidence from reassembly studies of other proteins [17–23], makes it increasingly clear that the specific association of discrete subdomains might be a general property of polytopic integral membrane proteins.

The TM topology of band 3 has been studied extensively by protein chemical and computation approaches (reviewed in [3,7,24,25]). Although the topology of the N-terminal eight spans is reasonably well established, the number and identity of the remaining spans that comprise the C-terminal portion remain controversial [8]. N-glycosylation is a co-translational event in which the luminal oligosaccharyltransferase adds the core oligosaccharide unit to an asparagine residue in the acceptor motif (Asn-Xaa-Ser/Thr), provided that it is more than 12 residues away from the end of the preceding TM span and further than 14 residues from the start of the following TM span [26]. The introduction of N-glycosylation motifs into putative intracellular (IC) and EC loops has proved to be useful in refining the topology of several multispanning membrane transporters, including cystic fibrosis TM conductance regulator [27], Glut1 hexose transporter [28] and sodium glucose co-transporter [29]. Recently, this approach has been used to refine the topology model of band 3 and to identify with increased accuracy the membranous limits of the TM spans [8]. However, it is notable that the insertion of a glycosylation site at Lys-743 (which is cytoplasmic in band 3 in red cells) was unexpectedly found to become partly N-glycosylated in microsomes [8]. This residue

Abbreviations used: BzNLT, benzoyl-Asn-Leu-Thr-N-methylamide; DNDS, 4,4'-dinitro-2,2'-stilbene disulphonate; EC, extracellular; ER, endoplasmic reticulum; IC, intracellular; GPA, glycophorin A; TM, transmembrane.

¹ To whom correspondence should be addressed (e-mail J.Groves@tesco.net).

might be exposed transiently to the lumen of the endoplasmic reticulum (ER) during folding, resulting in competition between the processes of folding and N-glycosylation. It is also possible that this region exhibits alternative TM topologies [30] as part of the functional cycle of the protein.

In the present study we have prepared complementary N- and C-terminal fragments that are separated in three previously undissected loops and show that two of these new pairs of fragments can generate transport-active protein when co-expressed in *Xenopus* oocytes. Secondly, we have prepared a series of constructs in which a short polypeptide containing a potential N-glycosylation acceptor site has been appended to the N-terminus of each C-terminal fragment. We compare the microsomal integration and orientation of the various TM spans of band 3 either individually or in the presence of one, several or all of the following spans. This approach has identified several spans that fold and insert broadly in accordance with the 'two-stage model' [31]. The topology of the other spans is influenced by interactions with adjacent regions. These results suggest that the membrane insertion of the TM spans of band 3 is not a simple sequential process involving successive signal anchor and stop transfer sequences [32]. These experiments allow us to define interacting TM regions in the protein, which provide constraints that should be helpful in building a three-dimensional structural model from the two-dimensional topology map of this polytopic membrane protein.

MATERIALS AND METHODS

Materials

The following were used: T7 mMessage mMachine[™] and T7 MEGascript[™] *in vitro* transcription kits (Ambion); rabbit reticulocyte lysate, canine pancreatic microsomes and TNT[™] coupled transcription and translation kit (Promega); Sculptor oligo-

nucleotide-directed mutagenesis kit, L-[³⁵S]methionine and Na³⁶Cl solution (Amersham); QIAprep plasmid mini-prep kit (Qiagen); restriction enzymes, Expand[™] proof-reading PCR system and T4 DNA ligase (Boehringer Mannheim); oligonucleotide synthesis performed by Professor L. Hall (University of Bristol, Bristol, U.K.); benzoyl-Asn-Leu-Thr-N-methylamide (BzNLT) peptide provided by Dr S. High (University of Manchester, Manchester, U.K.); other chemicals from Sigma, BDH or Fluka.

Band 3 cDNA constructs

The cDNA clones encoding human red-cell band 3 (pBSXG1.b3; residues 1–911), the membrane domain of band 3 (pBSXG1.b3mem; residues 361–911) and glycophorin A (pBSXG.GPA) have been detailed previously [6]. Asn-593 and Asn-642 were mutated individually to Ser by oligonucleotide-directed mutagenesis with the Sculptor kit, as described [33] to yield constructs pBSXG1.b3N593S and pBSXG1.b3N642S.

Construction and subcloning of band 3 membrane domain fragments

Ten complementary pairs of N- and C-terminal fragments of the band 3 membrane domain (b3mem) were prepared by truncation in each of the putative loops containing at least seven extra-membranous amino acid residues, i.e. not the loops between TM spans 2–3, 9–10, 11–12 and 13–14. B3mem fragment cDNA species (Table 1) were generated by PCR mutagenesis with pBSXG1.b3mem, pBSXG1.b3N593S or pBSXG1.b3N642S as template, as used previously [14,15]. The N-terminal fragments were prepared by the incorporation of a translation termination codon into the coding sequence of b3mem and deletion of the region of the cDNA on the 3' side of this termination codon.

Table 1 Construction of N- and C-terminal fragments of b3mem

The nomenclature and amino acid composition of each of the truncated fragments used in this paper are shown. The putative IC or EC loop location of each truncation site is shown in Figure 1. The site between Glu-755 and Val-756 might be located in either IC loop 5 or EC loop 6 owing to the unconfirmed topology of this region of band 3. An asterisk indicates the inclusion of the sequence Val-640 to Arg-646 (which contains the endogenous N-glycosylation site at Asn-642) in fragments truncated within EC loop 4. The PCR mutagenesis primers used to truncate the b3mem cDNA are shown; the *Nco*I or *Bst*EII sites used for cloning are underlined and the ATG initiator codon or TCA terminator anticodon is shown in bold.

N-terminal fragments				C-terminal fragments		
Truncated in loop	Fragment name	Amino acid residues	Anti-sense PCR mutagenesis primer	Fragment name	Amino acid residues	Sense PCR mutagenesis primer
EC1	bm(1)	M G ³⁶¹ –R ⁴³²	5'-GGCGGTAACCTCACC GGGTCTTTCTCCAG-3'	(2:14)	MG N ⁴³³ –V ⁹¹¹	5'-CCGCCATGGGTAACAGATGGGAGTGTC-3'
EC2	bm(1:3)	M G ³⁶¹ –T ⁴⁸¹	5'-GGCGGTAACCTCAGGTCTCGCAGAACGAGA-3'	(4:14)	MG N ⁴⁸² –V ⁹¹¹	5'-CCGCCATGGGTAACGGTCTAGAGTACATC-3'
IC2	bm(1:4)	M G ³⁶¹ –S ⁵¹⁷	5'-GGCGGTAACCTCAGGAGATGAAGCGGACCAG-3'	(5:14)	MG R ⁵¹⁸ –V ⁹¹¹	5'-CCGCCATGGGCCGCTATACCCAGGAGATC-3'
EC3	bm(1:5)	M G ³⁶¹ –Y ⁵⁵³	5'-GGCGGTAACCTCAATAAGTCTTCTGTAGTGGGT-3'	(6:14)	MG N ⁵⁵⁴ –V ⁹¹¹	5'-CCGCCATGGGTAACCTACAACGTGTGTATG-3'
IC3	bm(1:6)	M G ³⁶¹ –P ⁵⁹⁸	5'-GGCGGTAACCTCAAGGGAATAGGAGCTGTTCT-3'	(7:14)	M G ⁵⁹⁹ –V ⁹¹¹	5'-CCGCCATGGGTAAGTGCCTCGGGTCATC-3'
EC4	bm(1:7)	M G ³⁶¹ –K ⁶³⁹	5'-GGCGGTAACCTCACTTGAAGCCATCAGGCAC-3'	(*8:14)	M V ⁶⁴⁰ –V ⁹¹¹	5'-CCGCCATGGTGTCCAACCTCCTCAG-3'
EC4	bm(1:7*)	M G ³⁶¹ –R ⁶⁴⁶	5'-GGCGGTAACCTCACC GGGCTGAGGAGTTGG-3'	(8:14)	MG W ⁶⁴⁷ –V ⁹¹¹	5'-CCGCCATGGGCTGGGTCATCCAC-3'
IC4	bm(1:8)	M G ³⁶¹ –K ⁶⁹⁵	5'-GGCGGTAACCTCACTTGCCTCAGGTTTGCT-3'	(9:14)	M ⁶⁹⁶ –V ⁹¹¹	5'-CCGCCATGGTCAAGGGTCCGGC-3'
IC5/EC6	bm(1:10)	M G ³⁶¹ –E ⁷⁵⁵	5'-GGCGGTAACCTCACTCCTGGATCTGGGCTGC-3'	(11:14)	M V ⁷⁵⁶ –V ⁹¹¹	5'-CCGCCATGGTCAAGAGCAGCGGATC-3'
IC6	bm(1:12)	M G ³⁶¹ –Y ⁸²⁴	5'-GGCGGTAACCTCAGTAGGGCAGATCTGGGTG-3'	(13:14)	M V ⁸²⁵ –V ⁹¹¹	5'-CCGCCATGGTCAAGCGGGTGAAGACC-3'

Figure 1 Alternative topology models of b3mem, showing truncation positions in putative loops

(A) Diagram of the TM topology of the membrane domain of band 3 (b3mem) according to the traditional 14-span model [2,7,34]. Arrows indicate the approximate positions at which pairs of N- and C-terminal fragments were prepared in the putative IC or EC loops. The positions of Asn-593 (unglycosylated) and Asn-642 (glycosylated) are indicated by cross-hatched and solid filled circles respectively. The number of each TM span is indicated at its left. (B) Alternative topology proposed in [8]. The penultimate truncation site is within the region of unconfirmed topology.

Each anti-sense PCR primer contained a *Bst*EII restriction enzyme site and a termination codon, followed by the particular 17–20 nt sequence that perfectly matched the b3mem cDNA encoding the five to seven residues immediately before the truncation. The C-terminal fragments were prepared by insertion of a methionine codon and deletion of the region of the b3mem cDNA immediately upstream of this protein initiator codon. Each sense PCR primer contained an *Nco*I restriction enzyme site (underlined) and a methionine codon (in bold) followed by the particular 15–21 nt sequence that perfectly matched the portion of b3mem cDNA encoding the first five to seven residues that would follow in-frame with the initiator methionine residue. After amplification, the cDNA species were digested with *Bst*EII or *Nco*I, recircularized and cloned in *Escherichia coli* strain TG2. Putative IC and EC loops are indicated in Figure 1(A), in accordance with the 14-span model of band 3 [2,7,34]. An alternative topology map [8], in which the break between Glu-755 and Val-756 is assigned to EC loop 6 rather than IC loop 5, is shown in Figure 1(B).

The cloning of fragments bm(1:5), bm(1:7), bm(1:8), bm(1:12), (9:14) and (13:14) have been described [14,16]. Additional fragment cDNA species (not shown in Table 1) have been described in a previous study [15] with slightly different nomenclature: the C-terminal fragments designated gp(2:14), gp(4:14), gp(6:14) and gp(*8:14) have the the 19-residue cleav-

Table 2 Construction of small transmembrane fragments of b3mem

The nomenclature, amino acid composition and putative N- and C-terminal loop locations of each of the small internal b3mem fragments used in this paper are shown. An asterisk indicates the inclusion of the sequence Val-640 to Arg-646 (which contains the endogenous N-glycosylation site at Asn-642).

Putative loop location		Fragment name	Amino acid residues
N-terminus	C-terminus		
EC1	EC2	(2:3)	MG N ⁴³³ –T ⁴⁸¹
EC1	IC2	(2:4)	MG N ⁴³³ –S ⁵¹⁷
EC1	EC3	(2:5)	MG N ⁴³³ –Y ⁵⁵³
EC2	IC2	(4)	MG N ⁴⁸² –S ⁵¹⁷
EC2	EC3	(4:5)	MG N ⁴⁸² –Y ⁵⁵³
IC2	EC3	(5)	MG R ⁵¹⁸ –Y ⁵⁵³
EC3	IC3	(6)	MG N ⁵⁵⁴ –P ⁵⁹⁸
EC3	EC4	(6:7)	MG N ⁵⁵⁴ –K ⁶³⁹
EC3	EC4	(6:7*)	MG N ⁵⁵⁴ –R ⁶⁴⁶
IC3	EC4	(7)	M G ⁵⁹⁹ –K ⁶³⁹
IC3	EC4	(7*)	M G ⁵⁹⁹ –R ⁶⁴⁶
IC3	IC4	(7:8)	M G ⁵⁹⁹ –K ⁶⁹⁵
EC4	IC4	(*8)	M V ⁶⁴⁰ –K ⁶⁹⁵
EC4	IC4	(8)	M W ⁶⁴⁷ –K ⁶⁹⁵
EC4	IC5/EC6	(*8:10)	M V ⁶⁴⁰ –E ⁷⁵⁵
EC4	IC5/EC6	(8:10)	M W ⁶⁴⁷ –E ⁷⁵⁵
IC4	IC5/EC6	(9:10)	M ⁶⁹⁶ –E ⁷⁵⁵
IC4	IC6	(9:12)	M ⁶⁹⁶ –Y ⁸²⁴
IC5/EC6	IC6	(11:12)	M V ⁷⁵⁶ –Y ⁸²⁴

able signal sequence of glycophorin A (GPA) preceding the N-terminal residues of the fragments (2:14), (4:14), (6:14) and (*8:14). The cDNA coding sequence of each new construct was verified.

Addition of an N-terminal glycosylation marker to C-terminal fragments

A series of C-terminal fragments were prepared with a 22-residue sequence of the EC loop 4 of band 3 (which includes an N-glycosylation acceptor site) inserted before the normal N-terminus of the fragment. These 'Gtag' constructs encoded MSVSNSSARGWVIHPLGLRSES followed in-frame by the initiator methionine residue and the sequence of the corresponding C-terminal fragment. Equimolar quantities of the sense and anti-sense complementary oligonucleotides 5'-CATGAGC-GTTTCGAATTCCTCAGCCCGTGGTTGGGTAATTCAT-CCATTAGGGCTGCGTTCGGAGAG-3' and 5'-CATGCTC-TCCGAACGCAGCCCTAATGGATGAATTACCCAACCA-CGGGCTGAGGAATTCGAAACGCT-3' were annealed to form a short piece of double-stranded DNA with *Nco*I-compatible cohesive ends. Each C-terminal fragment cDNA (in pBSXG1) was linearized with *Nco*I and religated in the presence of a 200-fold molar excess of the annealed oligonucleotides. Ligation mixtures were then digested with *Nco*I to remove religated parental cDNA and subcloned in *E. coli*.

Construction of small fragment cDNA species without subcloning

cDNA species encoding small polypeptide fragments containing between one and four TM spans were prepared directly by PCR (Table 2) and the products were used directly to prepare cRNA. Both normal and 'Gtag' series of C-terminal fragment clones acted as templates for amplification with the various anti-sense primers that had been used to generate the N-terminal fragment series (Table 1). A unique sense PCR primer 5'-GGATGT-

GCTGCAAGGCGAT-3', which annealed to pBSXG1 [6] approx. 50 nt upstream of the T7 site, was used for all the amplifications.

Transcription *in vitro*, expression in *Xenopus* oocytes and cell-free translation

The T7 mMessage mMachine and T7 Megascript kits were used to prepare capped and uncapped cRNA species either by the transcription *in vitro* of *Hind*III-linearized plasmid DNA or directly from PCR-amplified DNA. The isolation of oocytes, microinjection of capped cRNA species and chloride uptake assays in the presence and absence of 2 mM 4,4'-dinitro-2,2'-stilbene disulphonate (DNDS) were performed as detailed previously [6,16].

Cell-free translation was performed in the rabbit reticulocyte lysate system with canine pancreatic microsomes and L-[³⁵S]-methionine for 60 min at 30 °C with either capped or uncapped cRNA species. Alternatively, a coupled transcription/translation system was used to translate closed circular plasmid cDNA. The competitive N-glycosylation acceptor peptide BzNLT was included in some reaction mixtures (60 µM final concentration), as described previously [33]. Microsomes were stripped of peripheral proteins and isolated by centrifugation in 7 mm × 20 mm tubes through 120 µl of alkaline cushion [250 mM sucrose/100 mM sodium carbonate (pH 11.5)] at 70 000 rev./min ($g_{av} = 189\,000$) for 20 min at 4 °C in a TLA100 ultracentrifuge (Beckman Instruments). Integral membrane protein translation products were separated on Laemmli [35] or tricine [36] SDS/PAGE gels; the efficiency of N-glycosylation was determined by scanning densitometry.

RESULTS

Nomenclature of C-terminal fragments

In this paper the nomenclature of fragments truncated in EC loops has been modified from that used previously [14–16,37]. First, the symbol '*' has been introduced to indicate the inclusion of the amino acid sequence Val-640 to Arg-646 (which includes the endogenous N-glycosylation site at Asn-642) in constructs truncated within EC loop 4. Thus fragments previously designated bm(1:7)_a and (8:14)_a are described here as bm(1:7) and (*8:14). Secondly, the C-terminal fragments previously designated (2:14), (4:14), (6:14) and (8:14)_a [15], which were all constructed with a GPA signal sequence preceding their N-terminus, are described here as gp(2:14), gp(4:14), gp(6:14) and gp(*8:14). This distinguishes them clearly from the leaderless C-terminal fragments (2:14), (4:14), (6:14) and (*8:14), which are introduced for the first time in this paper.

Functional co-expression of additional pairs of fragments divided in IC loops

By using PCR mutagenesis, we have prepared cDNA species encoding ten complementary pairs of N- and C-terminal band 3 membrane-domain (b3mem; residues 361–911 of band 3) fragments that together represent the entire amino acid sequence of b3mem (Table 1). Truncations were prepared in each putative IC and EC loop of b3mem that is considered to extend out of the membrane by at least seven residues [2,7,34]. However, our truncation site in IC loop 5 (between Glu-755 and Val-756) might have an alternative exofacial exposure at least some of the time during biosynthesis [8].

In an earlier paper [14] we showed that co-expressed complementary fragments of band 3 divided at positions in IC loop 4 (8+6 span combination) and IC loop 6 (12+2 spans) could

Figure 2 Chloride uptake into oocytes expressing new b3mem fragments divided in IC loops

Oocytes were co-injected with pairs of b3mem-fragment cRNA species at equimolar concentrations equivalent to 15 ng of b3mem cRNA, together with 2.5 ng of GPA cRNA. The amount of each cRNA injected was (per oocyte): 15.0 ng of b3mem, 5.7 ng of bm(1:4), 7.6 ng of bm(1:6), 9.9 ng of bm(1:8), 11.3 ng of bm(1:10), 12.2 ng of (5:14), 10.3 ng of (7:14), 8.0 ng of (9:14) and 6.6 ng of (11:14). Chloride influx (over 60 min) was measured 24 h after injection by using groups of 14–22 oocytes, either in Barths saline (hatched bars) or in Barths saline containing 2 mM DNDS (shaded bars). The DNDS-sensitive chloride influx can be derived from the difference between the mean values in the presence and the absence of DNDS. The bar indicates the S.E.M. of the chloride influx in each case. ***Highly significant ($P < 0.001$ in *t* tests) differences between the chloride influx in DNDS-treated and untreated oocytes and the DNDS-sensitive influx into parallel sets of injected and uninjected oocytes.

reassemble into the functional anion exchanger in *Xenopus* oocytes. In the present paper we have co-expressed complementary pairs of fragments divided in IC loop 2 (4+10 spans), IC loop 3 (6+8 spans) and IC loop 5/EC loop 6 (10+4 spans), which were not examined previously. Groups of *Xenopus* oocytes were co-injected with cRNA species encoding the pairs of fragments bm(1:4)+(5:14), bm(1:6)+(7:14) or bm(1:10)+(11:14). The cRNA encoding GPA was co-injected because GPA enhances the cell-surface expression of band 3 [6] and some complementary pairs of co-expressed band 3 fragments [14,15] in oocytes. After a 24 h incubation to allow only a limited amount of protein expression, the chloride uptake was measured in groups of 14–22 oocytes over a 1 h period in the presence and the absence of the band 3 inhibitor DNDS (Figure 2). Oocytes injected with b3mem cRNA mediated a high level of DNDS-sensitive chloride uptake, whereas oocytes co-injected with cRNA species encoding bm(1:8)+(9:14) mediated a slightly lower but nevertheless highly significant ($P < 0.001$) level of DNDS-sensitive uptake, as observed previously [14]. The oocytes co-expressing either bm(1:6)+(7:14) or bm(1:10)+(11:14) mediated a broadly similar level of DNDS-sensitive uptake ($P < 0.001$) as bm(1:8)+(9:14). In contrast, no DNDS-sensitive chloride uptake was detected in oocytes co-injected with the cRNA species encoding bm(1:4)+(5:14). In subsequent experiments, cRNA-injected oocytes were incubated for 48 h rather than 24 h to allow greater quantities of protein to be expressed before the assay of chloride uptake (results not shown). Under these conditions, the co-expressed fragments bm(1:6)+(7:14) and bm(1:10)+(11:14) mediated maximal levels of DNDS-sensitive influx ($P < 0.001$) similar to those of b3mem. However, injection of any of the fragment cRNA species individually, or co-injection of oocytes with the pair of cRNA species encoding

Figure 3 Chloride uptake into oocytes expressing new b3mem fragments divided in EC loops

Oocytes were injected with cRNA species at concentrations equimolar to those used for Figure 2. The amount of each cRNA injected was (per oocyte): 15.0 ng of b3mem, 3.6 ng of bm(1), 14.7 ng of gp(2:14), 14.2 ng of (2:14), 4.9 ng of bm(1:3), 13.6 ng of gp(4:14), 13.0 ng of (4:14), 6.6 ng of bm(1:5), 11.9 ng of gp(6:14), 11.3 ng of (6:14), 8.5 ng of bm(1:7), 9.9 ng of gp(*8:14), 9.3 ng of (*8:14) and 2.5 ng of GPA cRNA. Chloride influx was measured 24 h (A) or 48 h (B) after microinjection by using groups of 18–26 oocytes, as described in the legend to Figure 2. The variation in the DNDS-insensitive chloride uptake rate between experiments reflects a property of oocytes from different *Xenopus* toads. The bar indicates the S.E.M. of the chloride influx in each case. ***Highly significant ($P < 0.001$ in t tests) differences between the chloride influx in DNDS-treated and untreated oocytes and the DNDS-sensitive influx into parallel sets of injected and un.injected oocytes.

bm(1:4)+(5:14) did not result in any significant DNDS-sensitive uptake. We conclude that the integrity of IC loop 3, 4, 5 (Figure 2) or IC loop 6 [14] is not essential for functional reassembly from two fragments.

Functional co-expression of additional pairs of fragments divided in EC loops

In a previous paper [15] we showed that co-expressed complementary pairs of whole band 3 fragments truncated in EC loop 3 (5+9 span combination) or EC loop 4 (7+7 spans) would reassemble into the functional anion exchanger in *Xenopus* oocytes. In contrast, pairs of fragments truncated in EC loop 1 (1+13 spans) or EC loop 2 (3+11 spans) did not result in functional co-expression. In these experiments, each of the C-terminal fragments was preceded by the in-frame cleavable signal sequence of the type I single-spanning membrane protein GPA. This was employed to ensure that the initial TM span of the C-terminal fragments, here designated gp(2:14), gp(4:14), gp(6:14) and gp(*8:14), was incorporated into the membrane with the correct N_{out}–C_{in} orientation. For each of the fragments except gp(2:14), a high proportion of the GPA leader was cleaved by signal peptidase in microsomal membranes [15]. In this study we prepared the analogous C-terminal fragments without the GPA leader, here designated (2:14), (4:14), (6:14) and (*8:14), and tested whether these fragments would assemble into the correctly folded structure in oocytes (Figure 3).

Figure 3(A) shows the results of injecting oocytes with cRNA species encoding the fragments bm(1:5)+gp(6:14), bm(1:5)+(6:14), bm(1:7)+gp(*8:14) and bm(1:7)+(*8:14), either as complementary pairs or as individual fragment cRNA species. In each case, GPA cRNA was co-injected. After incubation for 24 h at 18 °C, all four pairs of fragments mediated a high level of DNDS-sensitive chloride uptake ($P < 0.001$), irrespectively of whether the C-terminal fragment possessed the GPA leader sequence. As expected, none of the individual fragments was functional. Figure 3(B) shows the results of a similar experiment

with the fragments bm(1)+gp(2:14), bm(1)+(2:14), bm(1:3)+gp(4:14) and bm(1:3)+(4:14). In this case, the influx assay was performed after a 48 h incubation of oocytes at 18 °C to improve the overall sensitivity of the assay to low levels of chloride uptake. However, none of the four complementary pairs of fragments showed statistically significant chloride uptake compared with un.injected oocytes in either the presence or the absence of DNDS. To confirm that the oocytes in the batch were competent, cRNA species encoding bm(1:5)+(6:14) were co-expressed and shown to mediate a very high ($P < 0.001$) level of DNDS-sensitive uptake. We conclude that the GPA leader sequence is neither essential for functional reassembly of the fragments truncated in EC loops 3 and 4 nor responsible for the lack of DNDS-sensitive chloride uptake by the fragments truncated in EC loops 1 and 2.

Orientation and integration of the N-terminal span of each C-terminal fragment

A key factor in determining the correct structure of C-terminal fragments is likely to be the orientation of the first TM span. In the 'sequential model' for insertion [32], the orientation of each span is determined by that of its immediately preceding span; in more complex models for assembly, proximal or distal spans, either preceding or following in the primary sequence, could assist in topogenesis and assembly, whether expressed in *cis*- or *trans*- polypeptides. We have used a glycosylation insertion mutagenesis approach to examine the TM orientation of the N-terminus of each C-terminal span. A 22-residue extension (the 'Gtag') was added to the N-terminal end of the C-terminal fragments (2:14), (4:14), and so on. The resulting series of constructs were designated Gtag(2:14), Gtag(4:14), and so on. The 'Gtag' sequence is identical with a large portion of EC loop 4 and contains an Asn residue in the context of the endogenous glycosylation site of b3mem. The distance between the newly inserted Asn residue in the 'Gtag' and the methionine that forms the initiator in normal C-terminal fragments was 18 residues.

Figure 4 Biosynthesis and N-glycosylation of C-terminal fragments in the cell-free system

(A) L - 35 S-methionine-labelled C-terminal fragments were translated in a rabbit reticulocyte lysate coupled transcription/cell-free translation system, with canine pancreatic microsomes. Plasmid DNA (400 ng in each 6.25 μ l reaction) was expressed in the presence or absence of 60 μ M competitive N-glycosylation acceptor tripeptide (BzNLT). Samples of translation mix (0.75 μ l) were analysed by SDS/PAGE and fluorography on 12% or 15% (w/v) Laemmli gels, as detailed in the Materials and methods section. Diglycosylated, monoglycosylated and unglycosylated products are indicated (+2, +1 and 0) for the outer lanes of each gel. (B) The percentage N-glycosylation was calculated from two independent scans of fluorographs; the bar represents the S.D. The horizontal broken line indicates the mean percentage N-glycosylation of b3mem at the endogenous Asn-642 from three independent reactions. Filled columns, diglycosylated product; hatched columns, monoglycosylated product.

This provided sufficient spacing from the putative membrane boundary to allow the oligosaccharyltransferase enzyme to gain access to the receptor site [26]. Constructs containing both the natural N-glycan addition site at Asn-642 and the 'Gtag' have the potential to add two N-glycan chains, as demonstrated with intact band 3 mutants [38]. Popov et al. [8] used a similar approach (inserting all 35 residues of EC loop 4) to assess the

topology of each of the smaller EC and IC loops of intact band 3. In principle, glycosylation insertion mutagenesis should detect either transient or stable exposure of the N-terminus to the ER lumen [8]. However, the possibility exists that the addition of a recombinant tag can affect the topology of the structural element to which it is appended and that the 'Gtag' could itself form interactions that favour certain configurations.

A series of coupled transcription/translation reactions were performed with the rabbit reticulocyte lysate TNT system (Promega) in the presence of canine pancreatic microsomes. To confirm the identity of glycosylated and unglycosylated bands, translations were performed with both 'Gtag' and normal C-terminal fragments in the presence and the absence of the competitive acceptor peptide BzNLT. The additional 22 residues caused each 'Gtag' fragment to run approx. 2 kDa higher than the corresponding normal C-terminal fragment. Fragments Gtag(2:14), Gtag(4:14), Gtag(5:14), Gtag(6:14) and Gtag(7:14) each contained the endogenous glycosylation site at Asn-642 as well as the potential glycosylation site in the 'Gtag'. In contrast, Gtag(9:14) and Gtag(11:14) contained only the glycosylation motif in the 'Gtag' sequence.

Figure 4(A) shows the results after separation by SDS/PAGE [12% or 15% (w/v) gel] and fluorography. As expected, inclusion of the BzNLT peptide in the translations caused the intensity of the slower-running glycosylated band(s) to decrease and the faster-running unglycosylated band to increase accordingly. The BzNLT-inhibitable bands ran at positions that were approx. 3 kDa (monoglycosylated) or 6 kDa (diglycosylated) higher than the corresponding unglycosylated band [identified in Figure 4(A) as the left-hand and right-hand lanes in each panel]. Fluorographs of gels derived from two independent cell-free translation experiments were quantified by scanning in the linear range of exposure of the film. The results (Figure 4B) are expressed as a percentage of total protein that was either monoglycosylated (hatched columns) or diglycosylated (filled columns). The glycosylation of each fragment was compared with that of the endogenous site in b3mem.

Fragment (2:14) was glycosylated at the endogenous site with a lower efficiency (32%) than either intact b3mem (70%) or the other C-terminal fragments (4:14), (5:14), (6:14) and (7:14) and (*8:14) (between 45% and 75% of the total protein). This suggests that the omission of TM span 1 alone affected the topology and/or structure of the loop between spans 7 and 8, at least in a proportion of the molecules. It is not clear whether this was due to inefficient membrane integration or a reversed topology for this loop. Furthermore, Gtag(2:14) contained a lower proportion of diglycosylated molecules than Gtag(4:14) and Gtag(6:14), which were predominantly translated with N-glycan chains added to both the 'Gtag' and the endogenous site. This indicated that Gtag(2:14) was less likely to insert with both its N-terminus and EC loop 4 facing the lumen than either Gtag(4:14) or Gtag(6:14).

Most of the Gtag(5:14) and Gtag(7:14) molecules were glycosylated at only one putative site. The level of this glycosylation (between 50% and 70%) was similar to that of (5:14), (7:14) and b3mem. The diglycosylation of only a minority (less than 10% of the total protein) background level of Gtag(5:14) and Gtag(7:14) molecules was consistent with the predicted cytoplasmic disposition of the N-terminus of these fragments. However, the fragment Gtag(7:14).N642S showed unexpectedly high levels of single-site glycosylation (40%). Because the only N-glycosylation motif in this fragment was located within the N-terminal 'Gtag', it seems that the N-terminal span of Gtag(7:14).N642S exhibited greater topological bivalence than that of Gtag(7:14). The presence of the endogenous glycan chain

in the loop between spans 7 and 8 might itself have stabilized the structure of this loop. This might have influenced the orientation of span 7 to favour the correct location of the N-terminus (in the cytoplasm).

Asn-593 is located in IC loop 3 within a putative glycosylation motif that is not modified in intact band 3 [33]. To confirm that Gtag(6:14) and (6:14) were glycosylated at Asn-642 and not at Asn-593, the corresponding fragments were constructed containing the N642S and N593S mutations. The fragments Gtag(6:14).N593S and (6:14).N593S yielded virtually identical results with those of Gtag(6:14) and (6:14) respectively; in contrast, Gtag(6:14).N642S and (6:14).N642S contained one fewer glycan chain than the wild-type fragments. These results confirmed that (6:14) and Gtag(6:14) were glycosylated in EC loop 4 and not in IC loop 3, similarly to intact band 3.

The constructs Gtag(9:14) and Gtag(11:14) contained only one glycosylation motif, located in the 'Gtag' sequence. However, translation of both these fragments resulted in substantial levels of glycosylation (65% and 46% respectively), indicating that in both cases the N-terminus was exposed to the ER lumen. Although the N-terminus of Gtag(11:14) is in a region of b3mem that has given controversial results in glycosylation scanning mutagenesis experiments with the whole molecule [8], the N-terminus of Gtag(9:14) is clearly predicted to be cytoplasmic in all topology models.

Interactions between TM spans affect topology and assembly

We prepared band 3 fragment cDNA species encoding fragments containing between one and four spans (Table 2), with and without the 'Gtag' sequence appended at the N-terminus. Uncapped cRNA species were translated; the microsomal membranes were purified by ultracentrifugation through an alkaline sodium carbonate cushion (Figure 5A). The 'Gtag' polypeptide sequence increased the apparent molecular mass of the polypeptide by approx. 2 kDa. Fluorographs of gels derived from up to four independent cell-free translation experiments were scanned and the combined results were expressed as a percentage of total protein that was either monoglycosylated or diglycosylated (Figure 5B).

TM spans 2–4

Fragments (2:3), (2:4), (2:5), (4) and (4:5) were each detected as a unique unglycosylated band (Figure 5A, lanes 1, 4, 7, 10 and 13), showing that these fragments could insert into the membrane in the absence of their preceding sequence. Although integrated into microsomes, the N-terminus of Gtag(2:3) was poorly glycosylated (less than 5%) (Figure 5A, lane 2) in comparison with Gtag(2:4) and Gtag(2:5), which were substantially glycosylated (46% and 55%; lanes 5 and 8). This indicated that the presence of span 4 increased the luminal exposure of the N-terminus of spans 2–3. Similarly, the N-terminus of Gtag(4) was poorly glycosylated (7%) (Figure 5A, lane 11) in comparison with Gtag(4:5) (38%; lane 14), indicating that the region containing span 5 facilitated the glycosylation of the N-terminus of span 4.

TM spans 5–6

Fragments (5), (6), (6:7) and (6:7*).N642S were all detected as single unglycosylated bands (Figure 5A, lanes 16, 19, 22 and 33). Fragment Gtag(5) was poorly glycosylated at the N-terminus (12%) (Figure 5A, lane 17), which was consistent with the predicted $N_{in}-C_{out}$ orientation of this span. Although we prepared the constructs Gtag(5:6) and Gtag(5:7), the chymotrypsin-

sensitive EC loop 3 (between spans 5 and 6) was too highly susceptible to proteolysis to show whether spans 6 and 7 affected the membrane integration of span 5 (results not shown). Fragments Gtag(6), Gtag(6:7) and Gtag(6:7*).N642S were each monoglycosylated at the N-terminus (65%, 86% and 85%) (Figure 5A, lanes 20, 23 and 34), which was consistent with the predicted $N_{out}-C_{in}$ topology of span 6. Fragment (6:7*), containing a C-terminal glycosylation site at Asn-642, was monoglycosylated at a similarly high level (65%) (Figure 5A, lane 27) to that of b3mem. The major product of Gtag(6:7*) and Gtag(6:7*).N593S was diglycosylated (64% and 64%) (Figure 5A, lanes 29 and 32) with smaller amounts of monoglycosylated fragment (17% and 19%; lanes 29 and 32). Fragments (6:7).N593S, Gtag(6:7).N593S, (6:7*).N593S and Gtag(6:7*).N593S (Figure 5A, lanes 25, 26, 31 and 32) all gave results identical with those obtained with the corresponding wild-type fragments (lanes 22, 23, 27 and 29), indicating that Asn-593 was not glycosylated in any of these constructs. Overall, these results indicated that the fragments containing spans 6–7 inserted into the membrane with both the N- and C-termini exposed to the ER lumen, as predicted by the topology model. The faint band observed in the 'Gtag' fragment samples at the same position as the corresponding unglycosylated wild-type fragment might have resulted from alternative initiation with the methionine codon downstream of the 'Gtag' sequence. We conclude that span 6 possesses particularly strong autonomous $N_{out}-C_{in}$ topogenic activity.

TM spans 7–8

Fragment (7) was observed as a unique unglycosylated band (Figure 5A, lane 36). Fragment (7*), which included the endogenous glycosylation site at Asn-642, was detected with an unexpectedly low level of C-terminal glycosylation (10%) (Figure 5A, lane 37). This contrasts with the result of fragments (6:7*) and (7:8), which were highly glycosylated (65% and 86%) (Figure 5A, lanes 27 and 43) at this site. Fragment Gtag(7*) contained putative glycosylation sites at both the N- and C-terminal ends of span 7. Although little diglycosylation was observed (3%) (Figure 5A, lane 41), Gtag(7*) was unexpectedly monoglycosylated to a greater extent (39%; lane 41) than the sum of Gtag(7) and (7*).

Fragments (7:8) and Gtag(7:8) were monoglycosylated to comparable extents (80% and 86%) (Figure 5A, lanes 45 and 43), and Gtag(7:8) was not diglycosylated (3%; lane 45). It is therefore highly likely that the monoglycosylation of Gtag(7:8) was at Asn-642 and the N-terminus of span 7 in these constructs was not exposed to the ER lumen, as predicted by the topology model. Although the major product of (7:8).N642S (Figure 5A, lane 47) was a band of similar mobility to that of unglycosylated (7:8), a minor product of lower molecular mass was observed uniquely in (7:8).N642S. This additional band, which was likely to have been a single-span proteolytic fragment, might indicate that EC loop 4 of (7:8).N642S was less tightly folded than (7:8) or Gtag(7:8).N642S. The N-terminus of Gtag(7:8).N642S was partly exposed to the ER lumen (15%) (Figure 5A, lane 48); although the level of glycosylation varied between samples, the mean was not as high as that of Gtag(7:14).N642S (Figure 4).

Fragments (*8) and (*8:10) were detected as monoglycosylated bands (38% and 61%) (Figure 5A, lanes 54 and 57), in contrast with (8) and (8:10) (lanes 53 and 56), which were unglycosylated. This showed span 8 to have autonomous $N_{out}-C_{in}$ topogenic activity, which was substantially increased in the presence of spans 9–10.

Figure 5 For legend see facing page.

TM spans 9–14

Fragments (9:10), (9:12), (11:12) and (13:14) were each detected as a unique unglycosylated band (Figure 5A, lanes 59, 62, 65 and 50), as expected. Whereas Gtag(9:10) showed little glycosylation of the N-terminus (8%) (Figure 5A, lane 60), which was consistent with the predicted $N_{in}-C_{out}$ orientation of this span, Gtag(9:12) was highly glycosylated at the N-terminus (64%; lane 63), as observed with Gtag(9:14) in Figure 4. This indicated that the interaction of spans 9–10 (which might comprise a single topogenic unit) with the span 11–12 region introduced a dynamic aspect or altered the structure of spans 9–10. Fragment Gtag(11:12) showed significant glycosylation (37%) (Figure 5A, lane 66) of the N-terminus of span 11, as observed with fragment Gtag(11:14) in Figure 4. This result agreed with the $N_{out}-C_{in}$ topology model for span 11–12 of band 3 [8]. In contrast, fragment Gtag(13:14) showed lower glycosylation (16%) (Figure 5A, lane 51) in agreement with the predicted $N_{in}-C_{out}$ orientation of span 13 in the topology models.

DISCUSSION

In this study we have continued our investigation of functional complementation by co-expressed pairs of N- and C-terminal band 3 fragments [14–16,37]. We have examined the effect of dividing band 3 in all the loops that have been predicted to extend from the bilayer by more than six residues (i.e. all except the loops between spans 2/3, 9/10, 11/12 and 13/14) and conclude that pairs of fragments truncated in any of six different putative loops within the C-terminal two-thirds of the molecule will co-assemble functionally from independent polypeptide chains. These pairs of N- and C-terminal fragments are folded correctly in the oocyte membrane and the integrity of the loops between spans 5/6, 6/7, 7/8, 8/9, 10/11 and 12/13 is not essential for anion exchange. Complementary pairs of fragments divided between spans 1/2, 3/4 and 4/5 were not functionally active but we have shown that at least two of these combinations co-assemble specifically in membranes [37].

Post-translational assembly of functional band 3 from two or more fragments requires that the biosynthesis and membrane integration of each polypeptide chain closely approximates that of the intact protein. Small-scale rearrangements of structure might accompany the interaction of fragment pairs within the bilayer but it is thermodynamically unlikely that large changes in structure will occur at this stage [31,39]. Although the factors influencing the topology and TM insertion of simple bitopic membrane proteins have been established, the features of eukaryotic multispanning membrane proteins that dictate their integration and assembly are less clear. According to the classic model, the N-terminus of the nascent polypeptide chain contains a hydrophobic signal sequence that interacts with the signal recognition particle to target the polypeptide to the membrane via the same Sec61p translocon complex as single-spanning

proteins; subsequent spans can then integrate sequentially with alternating stop transfer ($N_{out}-C_{in}$) and signal anchor ($N_{in}-C_{out}$) polarities and assemble in a co-ordinated manner (reviewed in [40]). However, the observation that C-terminal fragments of band 3 and other multispanning membrane proteins will fold tightly in the absence of the naturally preceding N-terminal residues indicates that these molecules possess multiple signals for insertion and co-translational assembly. In the present study we have compared the orientation of the first TM span of each C-terminal fragment with their topology in the current models of intact band 3, and investigated which spans can direct their own insertion or modify the insertion of adjacent TM spans. We discuss our results alongside those of Ota et al. [41,42], who recently assessed the topogenic functions of individual TM spans of band 3 in the cell-free system with the use of prolactin and/or *E. coli* leader peptidase fusion chimaeras.

TM spans 1–5

The region containing TM spans 1–5 will not permit division in any loops without a loss of function in oocytes (Figures 2 and 3B). This suggests that the fragment bm(1:5) corresponds to a subdomain of band 3 (see also [43]). Spans 1–5 are spaced in the primary sequence from the remainder of the membrane domain by the intervening EC loop 3 (Z-loop), which is variable in size and sequence across the anion-exchanger family [7]. Popov et al. [8] have shown that N-glycosylation acceptor sites introduced into EC loops 1 and 2 of band 3 are glycosylated with the same high efficiency as the endogenous glycosylation site in EC loop 4. This confirms the $N_{out}-C_{in}$ orientation of spans 2 and 4 in intact band 3. However, we observed that the C-terminal fragment Gtag(2:14) was less efficiently glycosylated at both its N-terminal and endogenous acceptor sites than was either Gtag(4:14) or Gtag(6:14) (Figure 4). Previously we found [15] that the glycoporphin leader sequence located at the N-terminus of fragment gp(2:14) was cleaved less efficiently by the luminal signal peptidase than that of gp(4:14) and gp(6:14). These results suggest that the absence of span 1 has a widespread effect on the topology and structure of (2:14). This effect is greater than the absence of spans 1–3 from (4:14). The glycosylation patterns of fragments Gtag(2:3), Gtag(2:4) and Gtag(2:5) demonstrate that span 4 is essential for the $N_{out}-C_{in}$ orientation of span 2, at least in the absence of span 1. Ota et al. [41] found that both span 2 and spans 2–3 had poor stop transfer ($N_{out}-C_{in}$) activity in the absence of span 1. However, they found good stop transfer activity of span 2 in constructs comprising spans 1–2 or 1–3. Furthermore, they showed that span 2 was dependent on the natural proximity of span 1 in the primary sequence: the introduction of a 200-residue spacer between the first two spans reduced by 3-fold the $N_{out}-C_{in}$ activity of span 2. Hamasaki et al. [44] have shown that spans 2–3 but not spans 1, 4 or 5 are released by the proteolysis of red-cell band 3 after partial denaturation with alkali; they concluded that spans 2–3 might interact with other parts of the protein rather than with the lipid.

Figure 5 Biosynthesis and N-glycosylation of single or groups of spans in the cell-free system

(A) cDNA species were amplified by PCR; cRNA transcripts were prepared directly. L-[³⁵S]methionine-labelled fragments were translated in the rabbit reticulocyte lysate cell-free system with canine pancreatic microsomes, as detailed in the Materials and methods section. Microsomal membranes were stripped of peripheral proteins and purified by ultracentrifugation (2 µl of translation mix) through an alkaline sucrose cushion (130 µl), then analysed by SDS/PAGE and fluorography on 12% or 14% (w/v) Tricine gels. Diglycosylated, monoglycosylated and unglycosylated 'Gtag' products (+2, +1 and 0) and unglycosylated product lacking the Gtag (←) are identified for the right-hand group of constructs of each gel. The positions of molecular mass markers are indicated (in kDa) at the left. (B) The mean percentage N-glycosylation of Gtag(2:3), Gtag(2:4), Gtag(2:5), Gtag(4), Gtag(4:5), Gtag(7), (7*), Gtag(7*), (7:8), Gtag(7:8), Gtag(7:8).N642S, Gtag(9:10), Gtag(9:12), Gtag(11:12), Gtag(13:14) and b3mem was determined by scanning fluorographs from four independent translations. Other values were derived from two independent translations. The bar on each column represents the S. D. The horizontal broken line indicates the mean percentage N-glycosylation at Asn-642 of b3mem. The constructs (2:3), (2:4), (2:5), (4), (4:5), (5), (6), (6:7), (6:7).N593S, (6:7*).N642S, (7), (7:8).N642S, (8), (8:10), (9:10), (9:12), (11:12) and (13:14) were unglycosylated and are not shown. Filled columns, diglycosylated product; hatched columns, monoglycosylated product.

Combining these results, we conclude that the correct membrane integration of spans 2–3 is dependent on interaction with both spans 1 and 4. South-East Asian hereditary ovalocytosis (SAO) is a genetic disease caused by the heterozygous presence of an abnormal band 3 that lacks nine residues (400–408) at the N-terminus of span 1 (reviewed in [45]). Although the TM spans of SAO band 3 retain their α -helical structure, the tertiary structure of the entire membrane domain is disrupted by the deletion of this portion of span 1 [46]. The integrity of the short loop between spans 1 and 2 seems to be important for the global structure of the entire molecule.

We found that the $N_{out}-C_{in}$ glycosylation of Gtag(4) was poor in comparison with that of Gtag(4:5) or Gtag(4:14); in contrast, Ota et al. [41] reported that span 4 had a good stop transfer activity individually. Our results suggest that there are interactions between the regions of TM spans 2 and 4, and between 4 and 5. It is conceivable that spans 4 and 5 interact with span 1 and spans 2–3, either inside or outside the translocon, to form the folded span 1–5 subdomain.

TM spans 6–8

In previous work, co-expression of the non-complementary fragments bm(1:5) and gp(*8:14) revealed that TM spans 6 and 7 are not essential for DNDS-sensitive chloride uptake [16]. The introduction and elimination of putative N-glycosylation sites in EC loops 3 and 4 has confirmed their luminal disposition in the cell-free system [8,38]. We show that span 6 or span 8 is inserted with an $N_{out}-C_{in}$ orientation in all the fragments that we have tested, independently of the influence of any neighbouring residues. The high level of glycosylation of the N-terminus of span 6 was particularly marked, in agreement with Ota et al. [42], who hypothesize a novel mechanism by which the integration of spans 4, 6 and 8 facilitates the insertion of their respective preceding TM segments.

In contrast, spans 5 and 7 showed little or no evidence of luminal exposure of the N-terminus, in agreement with the band 3 topology model. Interestingly, the EC C-terminus of fragment (7*) was very poorly translocated to the ER lumen, which is consistent with the suggestion that span 7 has inadequate translocation initiation ($N_{in}-C_{out}$) activity [41]. In contrast, we have found that Asn-642 of fragment (7:8) and the N-terminal fragment bm(1:8) are efficiently glycosylated in the cell-free system (Figure 5). This indicates that the translocation of EC loop 4 into the ER lumen is very efficient when span 7 is followed by span 8. However, we also show that fragment (6:7*) is glycosylated efficiently at the C-terminus (Figure 5); a construct similar to bm(1:7*) was reported to be N-glycosylated at Asn-642 at an intermediate level to that of intact b3mem [47]. These results demonstrate that span 6 facilitates the translocation of EC loop 4 into the lumen, even in the absence of span 8, and show that the interaction of spans 7 and 8 is not essential for efficient binding of the oligosaccharyltransferase. The earlier proposal of Tam et al. [47] that span 7 was an internal topogenic signal sequence was based on the efficient membrane integration and glycosylation of a fragment whose N-terminus was located at Thr-581; this fragment was 17 residues longer than our fragment (7*) and includes approx. 9 residues of putative span 6. We have found that fragment Gtag(7*) is more highly mono-glycosylated than either Gtag(7) or (7*) and suggest that the extended N-terminus of either Gtag(7*) or the span 7 construct of Tam et al. might have contributed to the correct insertion of the C-terminus of span 7 in the bilayer.

The N-terminus of Gtag(7:14).N642S was more highly mono-glycosylated than would have been predicted by the low level of

diglycosylation of Gtag(7:14) and the high level of mono-glycosylation of both Gtag(7:14) and (7:14) (Figure 4). Although less marked, a similar pattern was observed between Gtag(7:8).N642S, Gtag(7:8) and (7:8) (Figure 5). Thus the N-terminus of the N642S constructs is not fixed in the cytoplasm, as expected from its topology in the intact protein. Glycosylation at the endogenous site (Asn-642) might stabilize the native structure in this region: either the interaction of Asn-642 in the loop between spans 7 and 8 with the oligosaccharyltransferase (or with glycan-dependent chaperones such as calnexin) in the ER might stop the luminal translocation of the N-terminus of span 7, or alternatively the carbohydrate itself might form stable interactions with the polypeptide chain. By these mechanisms, an assemblage of individual spans, each contributing inherent topogenic properties of different strength, could become stabilized within the nascent protein. Glycosylation-incompetent band 3 was functionally expressed in oocytes but only when the chaperone GPA was co-expressed [33]. A possible role for glycosylation in protein folding has been suggested previously [48]. Lack of structure in an unglycosylated loop between spans 7/8 might account for the increased susceptibility of this locus to proteolysis in the fragment (7:8).N642S. Recent accessibility mutagenesis studies have suggested that this loop might fold into a vestibule in the mature protein [9].

TM spans 9–14

Topology models of band 3 differ extensively in their interpretation of the structure of the C-terminal end: the classic model contains six TM spans [2,7], whereas more recent models suggest four TM spans surrounding a region of undefined structure [8–11]. This alternative structure was proposed after potential N-glycosylation sites were introduced between Ala-735 and Ala-750 of intact band 3 and were unexpectedly found to be partly glycosylated [8]. This suggested that a region of the mature protein that had been established to be IC in red cells [49] was exposed to the ER lumen during biosynthesis. Hamasaki et al. [44] found that approx. 40 residues on the C-terminal side of span 9 (including putative span 10) could be released from the membrane of alkali-treated red cells by proteolysis *in situ*. Recently, spans 9–10 and 11–12 have been proposed to form single topogenic units with $N_{in}-C_{out}$ and $N_{out}-C_{in}$ orientations respectively [41], fully traversing the bilayer only twice.

We have shown previously that the C-terminal fragments (*8:14) and (11:14) are expressed stably in *Xenopus* oocytes in the absence of their complementary partners ([37], and J. D. Groves and M. J. A. Tanner, unpublished work). In contrast, fragment (9:14) could not be detected by immunoprecipitation from oocytes unless b3(1:8) was co-expressed [37]. Because (9:14) reassembles functionally with bm(1:8) in oocytes [14], these results indicate that the structure of a portion of fragment (9:14), probably spans 9–10 and is stabilized by interaction with span 8, even if this span is present in a *trans*-polypeptide. We show that the N-terminus of fragments Gtag(9:12) and Gtag(9:14) is exposed to the ER lumen but that this is not true of Gtag(9:10). Taken together, our results indicate that the $N_{in}-C_{out}$ orientation of span 9 might be strongly favoured by interaction with span 8 and less strongly favoured by interaction with spans 11–14; in addition, the N-terminus of span 9 (i.e. IC loop 4) might acquire a dynamic aspect in the presence of spans 11–12. These unexpected observations suggest that the topology of the span 9–10 region might not be fixed but might depend on the particular interactions with other parts of the band 3 molecule that predominate at a given time. This topological bivalence could reflect different interactions resulting

from changes in structure of this part of the protein that occur during the transport cycle. The N-terminus of fragments Gtag(11:12) and Gtag(11:14) was glycosylated, indicating exposure to the ER lumen similarly to that of the nearby region around Lys-743 in intact band 3 [8]. The structure–function and dynamics of the region containing spans 9–12 merit further detailed study.

Conclusion

The assembly of two or more regions of a membrane protein might be (1) directed by specific interactions between TM spans or loops that are distal in the primary sequence or (2) enforced by connectivity that physically prevents the separation of two adjacent elements of secondary structure. The ability of functional band 3 to co-assemble from pairs of fragments divided in any of six different loops indicates that connectivity is not essential in the assembly of the C-terminal two-thirds of the membrane domain. In contrast, co-expressed pairs of fragments divided in the loops between the N-terminal five spans were not functional. Whereas the correct interaction of spans 1 and 2 seems to depend on their close proximity in the primary sequence, spans 2 and 4 are more widely spaced in the polypeptide chain and hence are less likely to require connectivity for correct interaction. Because the membrane insertion of TM spans originates from the cytoplasmic side, then connectivity might be of greater importance in facilitating assembly across EC rather than an IC loops, in the absence of overriding topogenic factors (as with spans 6 and 8). Our results suggest that spans 1, 4–5, 6 and 8 might act as a scaffold for the assembly of spans 2–3, 7 and 9–10. The approach that we describe might provide a general method for dissecting the interactions between membrane-spanning regions in polytopic integral proteins.

We thank Professor Naotaka Hamasaki and Dr. Masao Sakaguchi for discussions and for communicating results before publication. This work was supported by a grant from the Wellcome Trust.

REFERENCES

- Jennings, M. L. (1989) *Annu. Rev. Biophys. Biophys. Chem.* **18**, 397–430
- Reithmeier, R. A. F. (1993) *Curr. Opin. Struct. Biol.* **3**, 515–523
- Tanner, M. J. A. (1997) *Mol. Membr. Biol.* **14**, 155–165
- Lepke, S. and Passow, H. (1976) *Biochim. Biophys. Acta* **455**, 353–370
- Lepke, S., Becker, A. and Passow, H. (1992) *Biochim. Biophys. Acta* **1106**, 13–16
- Groves, J. D. and Tanner, M. J. A. (1992) *J. Biol. Chem.* **267**, 22163–22170
- Wood, P. G. (1992) *Prog. Cell Res.* **2**, 325–338
- Popov, M., Tam, L. Y., Li, J. and Reithmeier, R. A. F. (1997) *J. Biol. Chem.* **272**, 18325–18332
- Tang, X.-B., Fujinaga, J., Kopito, R. and Casey, J. R. (1998) *J. Biol. Chem.* **273**, 22545–22553
- Fujinaga, J., Tang, X.-B. and Casey, J. R. (1999) *J. Biol. Chem.* **274**, 6626–6633
- Popov, M., Li, J. and Reithmeier, R. A. F. (1999) *Biochem. J.* **339**, 269–279
- Oikawa, K., Lieberman, D. M. and Reithmeier, R. A. F. (1985) *Biochemistry* **24**, 2843–2848
- Wang, D. N., Sarabia, V. E., Reithmeier, R. A. F. and Kuhlbrandt, W. (1994) *EMBO J.* **13**, 3230–3235
- Groves, J. D. and Tanner, M. J. A. (1995) *J. Biol. Chem.* **270**, 9097–9105
- Wang, L., Groves, J. D., Mawby, W. J. and Tanner, M. J. A. (1997) *J. Biol. Chem.* **272**, 10631–10638
- Groves, J. D., Wang, L. and Tanner, M. J. A. (1998) *FEBS Lett.* **433**, 223–227
- Kobilka, B. K., Kobilka, T. S., Daniel, K., Regan, J. W., Caron, M. G. and Lefkowitz, R. J. (1988) *Science* **240**, 1310–1316
- Bibi, E. and Kaback, H. R. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4325–4329
- Ridge, K. D., Lee, S. J. and Yao, L. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3204–3208
- Schmidt-Rose, T. and Jentsch, T. J. (1997) *J. Biol. Chem.* **272**, 20515–20521
- Wilkinson, B. M., Esnault, Y., Craven, R. A., Skiba, F., Fieschi, J., Kepes, F. and Stirling, C. J. (1997) *EMBO J.* **16**, 4549–4559
- Marti, T. (1998) *J. Biol. Chem.* **273**, 9312–9322
- Kohl, B., Wagner, C. A., Huelseweh, B., Busch, A. E. and Werner, A. (1998) *J. Physiol.* **508**, 341–350
- Tanner, M. J. A. (1989) *Methods Enzymol.* **173**, 423–431
- Reithmeier, R. A. F., Chan, S. L. and Popov, M. (1996) in *Handbook of Biological Physics*, vol 2 (Transport Processes in Eukaryotic and Prokaryotic Organisms) (Königs, W. N., Kaback, H. R. and Lolkema, J. S., eds.), pp. 281–309, Elsevier, New York
- Nilsson, I. and von Heijne, G. (1993) *J. Biol. Chem.* **268**, 5798–5801
- Chang, X. B., Hou, Y. X., Jensen, T. J. and Riordan, J. R. (1994) *J. Biol. Chem.* **269**, 18572–18575
- Hresko, R. C., Kruse, M., Strube, M. and Mueckler, M. (1994) *J. Biol. Chem.* **269**, 20482–20488
- Turk, E., Kerner, C. J., Lostao, M. P. and Wright, E. M. (1996) *J. Biol. Chem.* **271**, 1925–1934
- Skach, W. R., Calayag, M. C. and Lingappa, V. R. (1993) *J. Biol. Chem.* **268**, 6903–6908
- Popot, J.-L. and Engelman, D. M. (1990) *Biochemistry* **29**, 4031–4037
- Blobel, G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1496–1500
- Groves, J. D. and Tanner, M. J. A. (1994) *Mol. Membr. Biol.* **11**, 31–38
- Tanner, M. J. A. (1993) *Semin. Haematol.* **30**, 34–57
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Schägger, H. and von Jagow, G. (1987) *Anal. Biochem.* **166**, 368–379
- Groves, J. D., Wang, L. and Tanner, M. J. A. (1998) *Biochem. J.* **332**, 161–171
- Tam, L. Y., Landolt-Marticorena, C. and Reithmeier, R. A. F. (1996) *Biochem. J.* **318**, 645–648
- Lemmon, M. A. and Engelman, D. M. (1994) *Q. Rev. Biophys.* **27**, 157–218
- High, S. and Laird, V. (1997) *Trends Cell Biol.* **7**, 206–210
- Ota, K., Sakaguchi, M., Hamasaki, N. and Mihara, K. (1998) *J. Biol. Chem.* **273**, 28286–28291
- Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N. and Mihara, K. (1998) *Mol. Cell* **2**, 495–503
- Groves, J. D. and Tanner, M. J. A. (1999) *Biochem. J.* **344**, 699–711
- Hamasaki, N., Okubo, K., Kuma, H., Kang, D. and Yae, Y. (1997) *J. Biochem. (Tokyo)* **122**, 577–585
- Bruce, L. J. and Tanner, M. J. A. (1996) *Cell. Mol. Biol.* **42**, 953–973
- Moriyama, R., Ideguchi, H., Lombardo, C. R., Dort, H. M. V. and Low, P. S. (1992) *J. Biol. Chem.* **267**, 25792–25797
- Tam, L. Y., Loo, T. W., Clarke, D. M. and Reithmeier, R. A. F. (1994) *J. Biol. Chem.* **269**, 32542–32550
- Helenius, A. (1994) *Mol. Cell. Biol.* **5**, 253–265
- Jennings, M. L., Anderson, M. P. and Monaghan, R. (1986) *J. Biol. Chem.* **261**, 9002–9010

Received 28 May 1999/5 August 1999; accepted 27 September 1999