Identification of topological determinants in the N-terminal domain of transcription factor Nrf1 that control its orientation in the endoplasmic reticulum membrane

Yiguo ZHANG*† and John D. HAYES*†

*Biomedical Research Institute, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, Scotland, U.K., and †Laboratory of Cell Biochemistry and Gene Regulation, College of Bioengineering and Life Sciences, University of Chongqing, Chongqing 400044, China

Nrf1 [NF-E2 (nuclear factor-erythroid 2)-related factor 1] is a CNC (cap’n’collar) bZIP (basic-region leucine zipper) transcription factor that is tethered to ER (endoplasmic reticulum) and nucleus–vacuole membranes through its N-terminal signal peptide (residues 1–30). Besides the signal peptide, amino acids 31–90 of Nrf1 also negatively regulate the CNC-bZIP factor. In the present study we have tested the hypothesis that amino acids 31–90 of Nrf1, and the overlapping NHB2 (N-terminal homology box 2; residues 82–106), inhibit Nrf1 because they control its topology within membranes. This region contains three amphipathic α-helical regions comprising amino acids 31–50 [called the SAS (signal peptide-associated sequence)], 55–82 [called the CRACs (cholesterol-recognition amino acid consensus sequences)] and 89–106 (part of NHB2). We present experimental data showing that the signal peptide of Nrf1 contains a TM1 (transmembrane 1) region (residues 7–24) that is orientated across the ER membrane in an N\textsubscript{cyt}/C\textsubscript{lum} fashion with its N-terminus facing the cytoplasm and its C-terminus positioned in the lumen of the ER. Once Nrf1 is anchored to the ER membrane through TM1, the remaining portion of the N-terminal domain (NTD, residues 1–124) is transiently translocated into the ER lumen. Thereafter, Nrf1 adopts a topology in which the SAS is inserted into the membrane, the CRACs are probably repartitioned to the cytoplasmic side of the ER membrane, and NHB2 may serve as an anchor switch, either lying on the luminal surface of the ER or traversing the membrane with an N\textsubscript{cyt}/C\textsubscript{lum} orientation. Thus Nrf1 can adopt several topologies within membranes that are determined by its NTD.

Key words: detergent-resistant membrane, endoplasmic reticulum, NF-E2 (nuclear factor-erythroid 2)-related factor 1 (Nrf1), N-glycosylation, nuclear envelope, topology.

INTRODUCTION

Nrf1 [NF-E2 (nuclear factor-erythroid 2)-related factor 1], Nrf2 and Nrf3 are CNC (cap’n’collar) bZIP (basic-region leucine zipper) transcription factors [1–4]. All three CNC-bZIP proteins regulate cytoprotective genes that contain an ARE (antioxidant-response element, with minimal consensus sequence 5′-A\textsubscript{NG}(T/G)\textsubscript{A}/T\textsubscript{NG}nnC\textsubscript{NG}C\textsubscript{NG}'-3') in their regulatory regions [4–8]. Such genes include those encoding the glutamate–cysteine ligase catalytic and modifier subunits, class Alpha and Mu GST (glutathione transferase) isoenzymes, NAD(P)H:quinone oxidoreductase 1, haem oxygenase 1, ferritin and metallothionein [9–15]. Nrf1 and Nrf2 are widely expressed in mouse and human tissues [1–3,12,16]. By contrast, Nrf3 is particularly abundant in the placenta, and although it is present at modest levels in the liver, it is expressed to a much lesser extent in other tissues [4,17,18].

Gene-targeting experiments have revealed that Nrf1, Nrf2 and Nrf3 possess distinct biological functions. Global knockout of Nrf1 is embryonically lethal to the mouse, with the livers of foetal mice exhibiting increased levels of apoptosis, resulting from high levels of oxidative stress [19–21]. Furthermore, liver-specific disruption of Nrf1 in neonatal mice results in the development of non-alcoholic steatohepatitis and ultimately hepatoma [22]. Thus Nrf1 appears to be necessary for the expression of ARE-driven genes that are critical for liver development. Although Nrf1 is essential for foetal development, global knockout of Nrf2 in the mouse yields viable animals that develop normally [23]. However, Nrf2\textsuperscript{−/−} mice are intolerant of many environmental stressors, and are susceptible to acute and chronic injuries, inflammation and chemical carcinogenesis [24–28]. Lastly, Nrf3\textsuperscript{−/−} mice display no obvious phenotype [29], and it is possible that Nrf1 and/or Nrf2 can compensate the loss of Nrf3. Collectively, these studies indicate that among the CNC-bZIP factors, Nrf1 fulfils an indispensable and unique function in maintaining the redox status and integrity of the liver. This conclusion has prompted us to examine the activity of Nrf1 in greater detail.

Transcription factor Nrf1 exists as a number of isoforms that arise from alternative RNA splicing, translation from several internal ATG codons [2,6,16,30–32], and various post-synthetic modifications including glycosylation and possibly proteolytic cleavage [33,34]. We have reported previously that full-length Nrf1 is tethered to the ER (endoplasmic reticulum) through a non-cleavable N-terminal signal anchor sequence, and that it is glycosylated in this organelle through its NST (asparagine/serine/threonine-rich) domain [33]. Further studies have shown that Nrf1 is an integral membrane protein which can be sorted from the ER to the nuclear envelope [35]. Upon reaching the inner nuclear membrane, Nrf1 is presumably subject...
to retrotranslocation, whereupon it is recruited to the promoters of ARE-driven genes through its CNC-bZIP domain and gains access to the transcriptional machinery through its two acidic ADs (activation domains) (i.e. AD1 and AD2).

The NTD (N-terminal domain; residues 1–124) of Nrf1 was originally described as a region that negatively controls the CNC-bZIP protein [36]. In particular, it has been shown that progressive sequential deletion of residues 2–90 in Nrf1 produces a gradual incremental step-wise increase in its transactivation activity, suggesting that a substantial portion of the NTD contributes to its negative control. The NTD of Nrf1 contains an NHB1 (N-terminal homology box 1; residues 11–30) and an NHB2 (N-terminal homology box 2; residues 82–106), subdomains that were so named because they are also represented in the NTD of Nrf3 [8]. In Nrf1, NHB1 represents the signal anchor sequence [33], whereas in mouse Nrf3 it is cleaved [8]. In Nrf1, less is known about the function of NHB2 than NHB1. However, the NHB2 subdomain of Nrf3 is required for its activity as it contributes to the sorting of the factor within the ER [8]. It is therefore not known why the region between amino acids 31–90 of Nrf1 inhibits its activity, although it is probable that NHB2 contributes to repression as it overlaps with the sequence.

Knowledge of the topology of Nrf1 within the membrane is necessary in order to gain a better understanding of how the CNC-bZIP protein fulfils its biological functions. We have previously proposed that full-length Nrf1 possesses three TM (transmembrane) regions, TM1, TMi and TMc [35]. The predicted TM1 α-helix is formed by amino acids 7–24, and although it predominantly determines the topology of Nrf1 within the ER membrane, its orientation within the membrane is uncertain. In the present study we have determined the orientation of TM1 in ER membranes. We have also examined whether amino acids 31–90, as well as NHB2, might form amphipathic helices that contribute to the negative regulation of Nrf1.

**EXPERIMENTAL**

**Chemicals, antibodies and other reagents**

All reagents were of the highest quality available and were readily available commercially. The ER extraction kit and all chemicals were purchased from Sigma–Aldrich. PNGase F (peptide N-glycosidase F), Endo H (endoglycosidase H) and PK (Protease K) were obtained from New England Biolabs.

Rabbit polyclonal antibodies against CRT (calreticulin) and GFP (green fluorescent protein) were bought from Calbiochem and Abcam PLC respectively. A mouse monoclonal antibody against the V5 epitope and rabbit polyclonal antibodies against DsRed [a Discosoma sp. RFP (red fluorescent protein)] were from Invitrogen. Antisera against Nrf1 were produced (by Invitrogen) from COS-1 cells, as described previously [33,35]. These fractions were resuspended in 100 μl of 1 x isotonic buffer [10 mM Hepes (pH 7.8), containing 250 mM sucrose, 1 mM EGTA, 1 mM EDTA and 25 mM KCl]. Subsequently, membrane proteinase protection reactions were performed for 15, 30 or 60 min on ice in an aliquot (50 μg of protein) of the membrane-containing preparation with PK at a final concentration of 50 or 100 μg of protein/ml in either the presence or absence of 1 % (v/v) TX (Triton X-100). The reactions were stopped by incubation at 90°C for 10 min following the addition of 1 mM PMSF.

**Expression constructs**

Expression constructs for full-length mouse Nrf1 and its mutants, along with other expression constructs for mouse Nrf2, GFP and their chimaeric proteins fused with various portions of Nrf1, have been described previously [35,36]. These mutants and others used in the present study were created by PCR-directed deletion mutagenesis within the NTD of Nrf1, as described previously [37]. Two sandwiched fusion proteins, called DsRed/NTD/GFP and DsRed/N65/GFP, were engineered by inserting nucleotide sequences encoding the NTD of Nrf1 or its N-terminal 65 amino acids (N65) between the cDNAs for DsRed2 and GFP within the pDsRed2–GFP vector through either the Sall/KpnI or the HindIII/KpnI multiple cloning sites [35]. The fidelity of all cDNA products was confirmed by sequencing.

**Cell culture, transfection, luciferase reporter assays and statistical analysis**

Monkey kidney COS-1 cells (3 x 10^5) were seeded in a six-well plate and grown for 24 h in DMEM (Dulbecco’s modified Eagle’s medium). After reaching 70% confluence, the cells were transfected with a Lipofectamine™ 2000 (Invitrogen) mixture that contained an expression construct for wild-type Nrf1 or a mutant, along with either a P<sub>SV40</sub>GST2A-6×ARE-Luc (which contains six copies of the core ARE consensus sequence from rat GSTA2) or P<sub>SV40</sub>nqo1-ARE-Luc reporter plasmid (which contains only one copy of the ARE from mouse nqo1), together with pcDNA4/HisMax/lacZ; the latter encodes β-gal (β-galactosidase) and was used to control for transfection efficiency [35,36,38]. Luciferase activity was measured at approx. 36 h after transfection. The data were calculated as a fold-change (means ± S.D) of the activity obtained following transfection with an expression vector for the CNC-bZIP factor, when compared with that obtained following transfection with an empty pcDNA3.1/V5 His B vector. The data presented are each representative of at least three different independent experiments undertaken on separate occasions that were performed in triplicate. The significance of differences in the luciferase activity was determined using the Student’s t test and is shown as a P value.

**Immunocytochemistry and confocal microscopy**

These experiments were performed as described previously [33,36].

**Subcellular fractionation and membrane proteinase protection reactions**

The intact ER-rich membrane and nuclear fractions were prepared from COS-1 cells, as described previously [33,35]. These fractions were resuspended in 100 μl of 1 x isotonic buffer [10 mM Hepes (pH 7.8), containing 250 mM sucrose, 1 mM EGTA, 1 mM EDTA and 25 mM KCl]. Subsequently, membrane proteinase protection reactions were performed for 15, 30 or 60 min on ice in an aliquot (50 μg of protein) of the membrane-containing preparation with PK at a final concentration of 50 or 100 μg of protein/ml in either the presence or absence of 1 % (v/v) TX (Triton X-100). The reactions were stopped by incubation at 90°C for 10 min following the addition of 1 mM PMSF.

**Glycosylation mapping, deglycosylation reactions and Western blotting**

The Nrf1<sup>1–7N/Q</sup> mutant was constructed by site-directed mutation of all seven endogenous glycosylation consensus asparagine residues (Asn-X-Ser/Thr, where X is any amino acids except proline) into glutamine residues, thereby ensuring it would not be glycated in the ER lumen [35]. Using the cDNA for Nrf1<sup>1–7N/Q</sup> as a template, a series of N-linked glycosylation asparagine acceptor sites were introduced into its NTD. It was anticipated that, if the engineered glycosylation sites were translocated into the ER
lumen, the mutant Nrf1 protein would be glycated by in vivo addition of a glycan precursor Glc3Man9GlcNAc2; this technique is called glycosylation mapping mutagenesis [39]. Subsequently, this modification was detected using in vitro deglycosylation reactions, followed by Western blotting [33,35,40]. On some occasions, nitrocellulose membranes that had already been blotted with an antibody were washed for 30 min with stripping buffer [7 M guanidine hydrochloride, 50 mM glycine, 0.05 mM EDTA, 0.1 M KCl and 20 mM 2-mercaptoethanol (pH 10.8)] before being re-probed with additional primary antibodies [41]. The intensity of some Western blots was calculated using the Quantity One® software developed at the Bio-Rad Laboratories.

Bioinformatic analyses

The topology of Nrf1 around the membranes was predicted first using several bioinformatic algorithms, including the TopPred (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=toppred), TMpro (http://tpro.ihb.blcm.cs.cmu.edu/tmpro), HeliQuest (http://heliquest.ipmc.cnrs.fr/) and AmphipraSeek (http://npsa-pbil.ibcp.fr/cgi-bin/npsa-automat.pl?page=NPSA/npsa-amphipraseek.html) software programs. Thereafter, the predicted topologies were evaluated using molecular and biochemical experiments.

RESULTS

Identification of potential topological elements in the NTD of Nrf1

Several bioinformatic algorithms have predicted that the TM1 α-helix in Nrf1 is formed by amino acids 7–24 [35]. In the present study we have addressed the question of whether inhibition of Nrf1 activity by its NTD is due in part to the presence of other putative α-helical amphipathic regions within the domain, besides TM1, that are capable of interacting with membranes (Figure 1 and Supplementary Figure S1 at http://www.BiochemJ.org/bj/430/bj4300497add.htm). The first of these helices lies between amino acids 31 and 50, and we have designated it the SAS (signal peptide-associated sequence). It is apparent that SAS shares 55% similarity with the TM1 sequence, which suggests it may interact with membranes. Moreover, as shown in Figure 1(B), SAS also shows conservation with the core N-terminal TM sequence of Nvj1 (nucleus–vacuole junction protein 1), an integral membrane polypeptide anchored in the inner nuclear membrane of Saccharomyces cerevisiae [42]. The second predicted α-helical amphipathic region possesses a net basic charge and lies between amino acids 55 and 82. It overlaps with two potential CRACs (cholesterol-recognition amino acid consensus motifs), L/V-X_{i−5}−Y−X_{i−5}−R/K, which overlaps between amino acids 62 and 70, and also 74 and 82 (Figure 1C); we therefore called it the CRACs subdomain. The third possible α-helical amphipathic region has a net acid charge and lies between amino acids 89 and 106 (Supplementary Figure S1). As shown in Figure 1(D), it resides within NHB2. In addition, the NHB2 sequence shows 36% identity and 68% similarity with an N-terminal region of Arfgap1 [Arf (ADP-ribosylation factor)-GAP1 (GTPase-activating protein 1)], a membrane-bound protein that regulates membrane traffic and protein transport [43–45].

Consequences of loss of the SAS and CRACs subdomains on Nrf1 activity

We have previously reported that NHB1 negatively controls Nrf1 by targeting it to the ER membrane [33]. Similarly, we have reported that Nrf1 is also repressed by both NHB2 [35] and the region located between NHB1 and NHB2 that includes the SAS and CRACs motifs [36], but the reason why this subdomain within the NTD negatively controls the Cnc-bZIP factor is not known. To test the functional significance of SAS and CRACs, we transfected COS-1 cells with an expression construct for wild-type Nrf1 or Nrf1 mutants that lack either SAS and/or CRACs, along with a P_{svac}Gsta2-6×ARE-Luc reporter plasmid. Importantly, internal deletion of the SAS (residues 31–50), CRACs (residues 55–80) or both (residues 31–80) subdomains from Nrf1 gave rise to differential increases in the transactivation activity (Figure 2B). This result indicates that both the SAS and CRACs negatively regulate Nrf1.

Western blotting experiments were performed to determine whether the higher activity of Nrf1^55–80 is
Figure 2  Consequences of deletion of topological elements within the NTD of Nrf1

(A) Diagrammatic representation of the topological elements within NTD and their deletion mutants. (B) In the left-hand panel, an expression construct (1.2 μg of DNA) for wild-type Nrf1 or its indicated mutants, together with 0.6 μg of PSV40-GSTA2-β×ARE-Luc and 0.2 μg of the β-gal reporter plasmids, were cotransfected into COS-1 cells over 6 h, and thereafter the cells were transferred to fresh complete medium for an additional 24 h. Subsequently, cell lysates were prepared and luciferase reporter activity was measured. The data were calculated as a fold change (means ± S.D.) of the ARE-driven luciferase gene activity obtained following transfection of each of the Nrf1 expression constructs, when compared with reporter gene activity obtained following transfection of an empty pcDNA3.1/V5 His B vector; in all cases the background luciferase activity observed following transfection of an empty pGL3-promoter reporter gene plasmid and an empty pcDNA3.1/V5 His B vector was subtracted after it was first normalized for transfection efficiency on the basis of β-gal activity. The results presented are typical of those obtained from at least three independent experiments, each of which was performed in triplicate. The significance of the differences was determined by the unpaired Student’s t test and is shown as a P value (*P < 0.05, **P < 0.01 and ***P < 0.001) when compared with wild-type Nrf1 activity. In the right-hand panel, each of the above expression constructs (1.2 μg), along with 0.2 μg of the β-gal reporter plasmid, was cotransfected into COS-1 cells. Approx. 24 h after transfection, the cells were harvested in RIPA buffer and β-gal activity was measured. Following denaturation, the total cell lysates were (+), or were not (−), incubated for 1 h with 500 units of PNGase F, before they were resolved using 4–12 % SDS/NuPAGE, in Tris/Bis running buffer, and immunoblotted with antibodies against the V5 epitope. The amount of protein applied to each polyacrylamide sample well was adjusted to ensure equal loading of β-gal activity. The arrows indicate the position of Nrf1 isoforms estimated to be 120, 95 and 85 kDa. The molecular mass in kDa is indicated on the left-hand side of the gel. (C) COS-1 cells were cotransfected with 1.2 μg of an expression construct for wild-type Nrf1
accompanied by the appearance of novel forms of the factor. We have reported previously that wild-type Nrf1 migrates as both a non-glycated ~95 kDa band and a glycated ~120 kDa band during SDS/NuPAGE [33]. As shown in Figure 2(B) (right-hand panel), Nrf1 131−50 and Nrf1 45−80 both ran during SDS/NuPAGE as ~95 kDa and ~120 kDa bands; in both instances, digestion with PNGase F prior to electrophoresis eliminated the ~120 kDa band, thereby providing evidence that it was glycosylated. A non-glycated unstable ~85 kDa band was observed following electrophoresis of Nrf1 131−50, but not the Nrf1 45−80 mutant. These data indicate that the high activity of Nrf1 mutants lacking the SAS or CRACs subdomains is not associated with glycosylation status of the protein.

Although removal of the SAS subdomain in Nrf1 131−50 caused an obvious increase in its transactivation activity, confocal imaging revealed that this mutant protein was located primarily in the extranuclear ER and the nuclear envelope, but not within the nucleus (Figure 2C). By contrast, the Nrf1 45−80 mutant protein exhibited a relative increase in staining of the nucleus, but this was not accompanied by a decrease in its extranuclear staining when compared with wild-type Nrf1 (also see Supplementary Figures S2 and S3 at http://www.BiochemJ.org/bj/430/bj4300497add.htm).

The SAS and CRACs subdomains regulate the association of Nrf1 with membranes

Comparison of confocal images indicated that loss of SAS had no effect on the extranuclear localization of Nrf1 because both Nrf1 2−50 and Nrf1 2−30 mutant proteins gave similarly low levels (~20%) of extranuclear staining in COS-1 cells (Supplementary Figure S3). Rather, subcellular fractionation revealed an increased recovery of Nrf1 2−50 in the microsomal-rich fraction when compared with the recovery of Nrf1 2−30. The recovery of Nrf1 2−50 was either decreased by removal of CRACs, to give Nrf1 2−80, or diminished by deletion of both CRACs and NHB2, in the Nrf1 2−120 mutant (Supplementary Figure S4A at http://www.BiochemJ.org/bj/430/bj4300497add.htm). Subsequently, membrane proteasome protection reactions showed that a portion of Nrf1 2−30 was located on the cytoplasmic side of the intact ER/microsome-rich membranes, because this mutant protein was not protected by membranes during proteolysis (Supplementary Figure S4B). These data indicate that CRACs and NHB2, but not SAS, are required by Nrf1 to bind membranes.

SAS, CRACs and NHB2 in Nrf1 are transiently translocated into the ER lumen

The above experiments suggested that transactivation activity of Nrf1 may be modulated by its topology within membranes. To explore whether the NTD of Nrf1 resides within the lumen of the ER following its targeting to the organelle, we performed glycosylation-mapping mutagenesis using Nrf1 1−79/Q, a non-glycated mutant as the reference protein [35]. Nine novel glycosylation consensus sites were introduced at different indicated positions within the NTD of the Nrf1 1−79/Q mutant (Figure 3A). Among these engineered asparagine residues (eN), the two that were located closest to either the N-terminus or the C-terminus of TM1 (i.e. Nrf1 eN50 and Nrf1 eN230) were not glycosylated because no distinguishable change in their electrophoretic mobility was observed following PNGase F-catalysed deglycosylation (Figure 3B). This finding suggests that after ER targeting of Nrf1, TM1 is inserted within the ER membrane, but is not translocated into the lumen of this organelle. An alternative explanation is that, even if the C-terminus of TM1 were translocated into the lumen, it is associated so closely with the ER membrane that the eN230 residue is unable to reach the luminal active site of OST (oligosaccharyl transferase), a membrane-bound enzyme that catalyses the addition of a glycan precursor Glc3Man9GlcNAc2 to the amino group of the asparagine residues [39]. Similarly, the eN50 residue located in the N-terminus of SAS that adjoins TM1 was not glycated because the electrophoretic mobility of Nrf1 eN50 was not obviously altered following digestion of this mutant protein with either PNGase F or Endo H (Figures 3B and 3C).

Further glycosylation-scanning mutagenesis revealed that all of the eN residues that had been introduced at positions 51, 60, 78, 97, 112 and 125 in the NTD of Nrf1 were glycated in the ER lumen as their mobilities during NuPAGE differed following PNGase F digestion (Figure 3B). These observations indicate that amino acids 51–125 around and within the SAS, CRACs and NHB2 are translocated transiently into the ER lumen. However, close examination of their electrophoretic mobilities revealed that the Nrf1 eN97 and Nrf1 eN120 mutant glycoproteins were deglycosylated by Endo H (Figure 3C), but they were not extensively deglycosylated by PNGase F as the latter digestion did not alter their migration during NuPAGE (Figure 3B). These two different glycosidase reactions suggest that amino acids 97 and 112 located within and around NHB2 are probably embedded within a tightly folded conformation.

ARE-driven reporter luciferase assays showed a marked increase in transactivation by glycated Nrf1 eN51, in which eN51 was located in a possible loop between SAS and CRACs (Figure 3D). Other glycated proteins Nrf1 eN60, Nrf1 eN78 and Nrf1 eN120 were translocated transiently into the ER lumen. However, close examination of their electrophoretic mobilities revealed that the Nrf1 eN97 and Nrf1 eN120 mutant glycoproteins were deglycosylated by Endo H (Figure 3C), but they were not extensively deglycosylated by PNGase F as the latter digestion did not alter their migration during NuPAGE (Figure 3B). These two different glycosidase reactions suggest that amino acids 97 and 112 located within and around NHB2 are probably embedded within a tightly folded conformation.

The Authors Journal compilation © 2010 Biochemical Society
Figure 3  The SAS, CRACs and NH2 subdomains are transiently translocated into the ER lumen

(A) The upper panel shows the structural subdomains in the NTD of Nrf11–7NQ mutant, in which seven endogenous glycosylation consensus sites in Nrf1 had been removed by mutation of the relevant asparagine residues into glutamine residues [35]. The lower amino acid sequence shows the positions of eight new glycosylation sites that were introduced into the NTD of Nrf11–7NQ protein. The new engineered asparagine (eN) glycosylation sites are represented as white letters on a black background. (B) The constructs for Nrf1 and its asparagine mutants were cotransfected, along with the β-gal reporter plasmid, into COS-1 cells. Approx. 24 h after transfection, cell lysates (30 μg of protein) were prepared and subjected to in vitro deglycosylation using 500 units of PNGase F for 1.5 h at 37°C. The resulting samples were resolved using SDS/NuPAGE containing 4–12% polyacrylamide in Tris/Bis buffer, followed by Western blotting. (C) The above protein samples were also subjected to a deglycosylation reaction with 500 units of Endo H for 1.5 h at 37°C. The digests were resolved in a 7% SDS/NuPAGE Tris/acetate gel and Nrf1 protein was identified by Western blotting. For (B and C), the molecular mass in kDa is indicated on the left-hand side of the gel. (D) COS-1 cells were cotransfected with each of the above expression constructs together with P2ARE GSTA2-6×ARE-Luc and β-gal reporter plasmids, before luciferase reporter assays were carried out 24 h later, as described in Figure 2.
**Figure 4**  The TM1 region of Nrf1 is inserted with an N<sub>cyt</sub>/Clum orientation within ER membranes

(A) Shows amino acids 1–50 of wild-type Nrf1 and the point mutations that were introduced to create potential GAA and ASA cleavage sites for SPase.  
(B) Total lysates prepared from COS-1 cells that had been transfected with an expression construct for Nrf1, Nrf1<sup>L12/13A</sup> or Nrf1<sup>L19/21A</sup> were subjected to in vitro PNGase F deglycosylation reactions, followed by Western blotting with anti-V5 antibodies.  
(C) Total lysates from COS-1 cells that had been transfected with an expression construct for NTD/GFP<sup>×2</sup>, N65/GFP<sup>×2</sup> or their point mutants were resolved by 7% SDS/NuPAGE using the Tris/acetate running gel buffer system, before being visualized by immunoblotting with anti-GFP antibodies.  
(D) An expression construct (1.2 μg of DNA) for Nrf1, Nrf1<sup>L12/13A</sup> or Nrf1<sup>L19/21A</sup>, together with 0.6 μg of Pkn<sub>mc</sub> ARE-Luc and 0.2 μg of the β-gal reporter plasmids, was transfected into COS-1 cells over 6 h, and then the cells were transferred to fresh complete medium for an additional 24 h. Subsequently, luciferase reporter activity was measured. The data were calculated as a fold change (means ± S.D.) of the activity observed following transfection with the reporter plasmid and an empty pcDNA3.1/V5 His B expression vector. The results presented are typical of those obtained from at least three independent experiments, each of which was performed in triplicate. The significance of the differences was determined using the unpaired Student’s t-test and is shown as a P value.  
(E) COS-1 cells were transfected with 1.2 μg of each of the above expression constructs for Nrf1 and its mutants, along with 0.6 μg of an ER/DsRed construct. Approx. 24 h following transfection, the subcellular location of Nrf1 was examined by immunocytochemistry, followed by confocal imaging as described in Figure 2(C), with the exception that nuclear DNA was stained blue by DAPI. Scale bars = 20 μm.

To detect by SDS/NuPAGE the loss of only 13 or 21 amino acid residues from full-length Nrf1 proteins (Figure 4B). To resolve this problem, we further examined the electrophoretic properties of N65/GFP<sup>×2</sup> and NTD/GFP<sup>×2</sup> fusion proteins and their point mutants because they were considerably smaller than Nrf1 and did not contain a functional glycosylation consensus site. Immunoblots of COS-1 cell lysates expressing these fusion proteins revealed that both N65/GFP<sup>×2</sup> and NTD/GFP<sup>×2</sup> were clearly resolved from GFP<sup>×2</sup> by SDS/NuPAGE (Figure 4C); during electrophoresis, N65/GFP<sup>×2</sup> and NTD/GFP<sup>×2</sup> each gave...
a clear single band with estimated molecular masses of 65 kDa and 73 kDa respectively, whereas GFP × 2 migrated with an apparent mass of 58 kDa. The relative mobilities of N65/GFP × 2 and NTD/GFP × 2 were consistent with them containing, respectively, 65 and 125 amino acids more than GFP × 2. Importantly, a distinct cleaved polypeptide of ∼62.5 kDa was detected in cell lysates expressing N65L19/GFP × 2 (Figure 4C). Similarly, a cleaved polypeptide of ∼71 kDa with a mobility similar to the full-length ∼73 kDa protein was also observed in the cells expressing NTDL19/GFP × 2. These findings demonstrate that the pentapeptide Ala19-Ser-Ala-Ile-Gly23 in N56L19/GFP × 2 and in NTDL19/GFP × 2 is located around the membrane luminal interface in close vicinity to the active site of SPase. By contrast, neither N65L12/GFP × 2 nor NTDL12/GFP × 2 yielded a cleaved polypeptide, suggesting that the Gly11-Ala-Ala13 tripeptide is positioned in the membrane hydrocarbon centre or close to the membrane cytoplasmic interface. Therefore we deduce that the TM1 α-helix is inserted in membranes with an orientation of Ncyt/C lum (i.e. with its N-terminus in the cytoplasm and its C-terminus within the ER lumen).

Although the Nrf1L19/GFP mutant migrated primarily as a 120 kDa glycoprotein, it also yielded three shorter glycopolypeptides with molecular masses of ∼36 kDa, ∼52 kDa and ∼65 kDa (Figure 4B). However, the appearance of these additional Nrf1L19/GFP products did not result in obvious changes in either its transcription activating activity (Figure 4D) or subcellular distribution (Figure 4E) when compared with wild-type Nrf1. The finding that Nrf1L19/GFP yields multiple forms is consistent with amino acids at positions either 21–22 or 23–24 being proteolytically cleaved by a luminal SPase, followed possibly by sequential proteolytic cleavage around and within its NST domain. By contrast, the Nrf1L12/GFP mutant was expressed exclusively as a nuclear non-glycated 95 kDa protein (Figures 4B and 4E), and exhibited increased activity (Figure 4D). This result indicates that the leucine residues 12 and 13 in the TM1 region are situated in the hydrocarbon centre of the membrane lipid bilayer and their hydrophobic properties are essential for insertion of Nrf1 into the ER membranes.

No proteolytical cleavage occurs within the N-terminal 170 amino acids of Nrf1

To further characterize the effect of the Ncyt/C lum orientation of TM1 on its post-insertion processes, we created three chimaeric proteins, in which the NTD of Nrf1, the N-terminal 156 amino acids of Nrf1 (N156), and the N-terminal 170 amino acids of Nrf1 (N170) were attached to the N-terminus of Nrf2, giving NTD–Nrf2, N156–Nrf2 and N170–Nrf2. Western blotting of total cell lysates expressing these chimaeric proteins showed that they gave an apparent single electrophoretic band and its migration in the NuPAGE gel increased following PNGase F digestion (Figure 5). No cleaved polypeptide was found in these chimaeric proteins. These results suggest that the TM1-containing signal peptide of Nrf1 anchored the Nrf2-containing chimaeric proteins to the ER, where their remaining C-terminal portions were translocated into the lumen and glycosylated. We conclude that the NTD, N156 and N170 portions of the fusion proteins are not proteolytically cleaved by an SPase, or other proteases, under normal homeostatic conditions.

The TM1-connecting portions in the NTD of Nrf1 are repositioned from the ER lumen into the extraluminal compartments

Our experiments above suggested that during the topogenesis of Nrf1 its TM1 region is inserted with an Ncyt/C lum orientation within ER membranes, and its associated SAS, CRACs and NHB2 subdomains are translocated into the lumen of the ER. However, it is not known whether any portion of Nrf1 that lies C-terminal to TM1 is repositioned into the membrane cytoplasmic side of the ER. To address this question, we performed time-lapse membrane PK protection reactions with the intact ER and nuclear fractions purified from COS-1 cells that had been transfected with an expression construct for either DsRed/NTD/GFP or DsRed/N65/GFP. Immunoblotting using either anti-DsRed or anti-GFP antibodies revealed that the full-length DsRed/NTD/GFP sandwiched fusion protein of ∼68 kDa was almost totally digested following incubation of the intact ER fractions with PK for 30 min or 60 min (Figure 6A). By contrast, the nuclear ∼68 kDa DsRed/NTD/GFP fusion protein appeared to be partially digested by the protease (Figure 6A). A cleaved polypeptide ladder of between 10 kDa and 60 kDa was detected by immunoblotting, that reacted strongly with anti-GFP antibodies (Figure 6A, middle panel) and weakly with anti-DsRed antibodies (upper panel). Two PK-digested bands of ∼32 kDa and ∼45 kDa cross-reacted with antibodies against GFP (middle panel) rather than DsRed (upper panel). Based on their molecular masses, the immunoreactive bands appear to represent two small GFP-fusion proteins of ∼32 kDa and ∼45 kDa that contain a short C-terminal segment of the NTD of Nrf1. Further examination showed that, as the proteinase reaction time was extended from 5 min to 60 min, the abundance of the ∼32 kDa and ∼45 kDa polypeptides gradually decreased to a low level, but they were not completely abolished, even in the presence of TX. A similar time-course response for the ∼32 kDa polypeptide was observed in the nuclear membrane PK digestions. Interestingly, no band of between ∼33 kDa and ∼44 kDa was observed in the nuclear membranes. These data suggest that, during topogenesis of Nrf1, TM1 is orientated in an Ncyt/C lum fashion within membranes and SAS, CRACs and NHB2 are reintegrated and repositioned from the lumen into extraluminal compartments around the ER and nuclear envelope membranes (Figure 6C, left-hand side).

Similar proteinase protection of ectopic DsRed/N65/GFP against PK digestion revealed that a full-length fusion polypeptide of ∼62 kDa was recovered in the ER and nuclear fractions (Figure 6B). Incubation of the ER fraction with PK for 5–30 min resulted in loss of most of the ∼62 kDa DsRed/N65/GFP.
Figure 6 The extra-TM1 regions of NTD are reintegrated from the lumen into the extraluminal environment

COS-1 cells were transfected with an expression construct for either DsRed/NTD/GFP (A) or DsRed/N65/GFP (B), two sandwiched fusion proteins containing either the NTD or the N65 flanked N-terminally by DsRed and C-terminally by GFP. The intact ER and whole nucleus (WN) fractions were purified and equal amounts (50 μg of protein in 40 μL of 1× isotonic buffer) were examined. Membrane protease protection reactions were carried out by incubating (on ice) either ER or WN fractions with 100 μg/ml (+) or absence (−) of 1% TX for 5, 15, 30 and 60 min. The reaction products were resolved in a 4–12% SDS/NuPAGE Tris/Bis gel and identified by Western blotting with rabbit polyclonal antibodies against DsRed (upper panel) or GFP (middle panel). The anti-DsRed antibody-blotted nitrocellulose membrane was stripped and re-probed with antibodies against GFP and CRT. For (A and B), the molecular mass in kDa is indicated on the left-hand side. (C) The cartoon shows two proposed topologies of DsRed/NTD/GFP and DsRed/N65/GFP. The scissors indicate potential PK cleavage sites that are not protected by membranes.

protein. In a 60 min digest, PK almost completely eliminated the ∼62 kDa fusion protein in the ER, but not in the nuclear fraction. An apparent PK-digested band of ∼32 kDa cross-reacted with antibodies against GFP, but not against DsRed, and its abundance was modestly increased when the enzyme reaction time was extended. An additional weak PK-digested band of ∼38 kDa was also detected by immunoblotting with anti-GFP antibodies (middle panel). When compared with the ∼28 kDa GFP protein, the relative mobilities of the PK-digested ∼32 kDa and ∼38 kDa bands are consistent with their identification as two GFP-fusion polypeptides that contain the entire N65 amino acids or its C-terminal portion respectively. These data indicate that, after the N-terminal 65 amino acids of Nrf1 are integrated in the membrane through TM1, its flanking N-terminal portion is partitioned in the cytoplasmic or nucleoplasmic sides, and its adjacent SAS subdomain can be repartitioned from the lumen into the extraluminal environments closely associated with the membranes (Figure 6C, right-hand side). Furthermore, these two PK-digested polypeptides of ∼32 kDa and ∼38 kDa were not completely destroyed after the ER membrane was disrupted by 1% TX, suggesting that the N-terminal 65 amino acid portion of Nrf1 may diffuse from the TX DRM (detergent-resistant membrane) microdomain to the detergent-sensitive membrane regions.

SAS, but not CRACs, is required to repartition Nrf1 within the ER

To identify roles for SAS and CRACs in determining the topology of full-length Nrf1 protein, we transfected COS-1 cells with expression constructs for wild-type Nrf1 and three internal deletion mutants of the CNC-bZIP protein, followed by subcellular fractionation and time-lapse membrane
Figure 7 Repartitioning of the NTD of Nrf1 is modulated by SAS, CRACs and NHB2 subdomains

COS-1 cells were grown in a 100-mm dish and transfected with 6 μg of an expression construct for wild-type Nrf1 (A), Nrf1\(\Delta^{31-50}\) (B), Nrf1\(\Delta^{55-80}\) (C) or Nrf1\(\Delta^{96-106}\) (D). Approx. 24 h following transfection, the cells were subjected to subcellular fractionation, followed immediately by membrane proteinase protection reactions that were carried out for 5, 15, 30 or 60 min in 20 μl of 1× isotonic buffer containing 100 μg/ml (+) or not (−), together with 100 μg of proteins from the intact ER or whole nuclei (WN) fractions. The reaction products (10 μg of proteins diluted in 10 μl of 1× loading buffer) were resolved in a 4–12% SDS/NuPAGE Tris/Bis gel and Nrf1 protein was identified by Western blotting with anti-V5 antibodies. The antibody-blotted nitrocellulose membrane was stripped and re-probed with antibodies against CRT. The intensity of these Western blots was calculated using the Quantity One® software. The relative amount of Nrf1 after PK digestion was estimated by dividing its immunoreactive band intensity with that obtained for CRT, and then the resulting values were normalized to the total non-digested amounts and shown as percentages at the bottom. In addition, these results are shown graphically in Supplementary Figure S5 at http://www.BiochemJ.org/bj/430/bj4300497add.htm.

PK protection reactions. Figure 7(A) shows that, following incubation of ER membranes with PK for 5–60 min, the abundance of the ∼120 kDa wild-type Nrf1 glycoprotein was reduced by proteolytic digestion from ∼60% to ∼10% of the level observed following incubation in the absence of PK (see Supplementary Figure SSA at http://www.BiochemJ.org/bj/430/bj4300497add.htm). The nuclear membrane protease protection reactions with PK caused between ∼75% and 85% of the ∼120 kDa Nrf1 glycoprotein to be digested. Upon addition of 1% TX into these membrane reactions, all remaining Nrf1 was destroyed by PK; it should be noted that no PK-digested polypeptides of between 6 kDa and 94 kDa were observed by immunoblotting with antibodies against either Nrf1 or its C-terminal V5 tag (also see Supplementary Figure S6 at http://www.BiochemJ.org/bj/430/bj4300497add.htm). By contrast, the ER luminal protein CRT was protected by the membranes from PK digestion (Figure 7A). Taken together, these data suggest that a large C-terminal portion of Nrf1 is dynamically retrotranslocated from the ER lumen to the cytoplasmic and/or nucleoplasmic sides. In addition, all Nrf1 isoforms of less than 120 kDa were recovered primarily in the nuclear fractions and rapidly destroyed by digestion with PK (Figure 7A), suggesting that these products resided in the extraluminal compartments and thus are not protected by membranes.

When compared with wild-type Nrf1, the Nrf1\(\Delta^{31-50}\) mutant protein was preferentially retained in the ER lumen following PK digestion (Figure 7B). In the intact ER membrane, ∼50% of Nrf1\(\Delta^{31-50}\) resisted PK digestion for 15–60 min. In the nuclear membrane PK reactions, digestion of Nrf1\(\Delta^{31-50}\) did not differ significantly from the wild-type protein. Similar results were obtained from the membrane PK protection reactions with the Nrf1\(\Delta^{55-80}\) mutant protein (Supplementary Figures S5A and S5C). Further experiments showed that PK digestion of the Nrf1\(\Delta^{55-80}\) mutant, lacking CRACs alone (Figure 7C), gave a similar result to those obtained for the wild-type Nrf1 (Supplementary...
Mechanisms that control the topology of Nrf1 within membranes

Our results demonstrate that Nrf1 exists in several topologies that enable it to partition into different compartments around membranes ([35] and the references therein). We propose that the different topologies of Nrf1 could arise through co-translational and/or post-translational mechanisms. According to our present understanding [50–53], the initial topology of Nrf1 is determined by the amino acid composition of the nascent polypeptide as it is synthesized. This CNC-bZIP protein contains several topogenic signals (e.g., hydrophobic segments and the positive/negative charge distribution of associated residues) and glycosylation recognition motifs (e.g., Asn-X-Ser/Thr) in its primary structure. In ribosome-polypeptide-translocon Sec61 complexes, the ribosome serves as a platform for the cotranslational processing of newly synthesized Nrf1 polypeptides, whereas translocon Sec61 provides a permissive environment required for the concurrent insertion of hydrophobic regions of Nrf1 into the rough ER membranes while simultaneously positioning hydrophilic amino acid flanking regions on either side of the membrane according to the positive-inside and/or charge difference rules. We therefore envisage that the NH1 signal peptide of Nrf1 directs its TM1 to enter the ER membrane-bounded translocon Sec61 complex, whereupon it is inserted into membranes with an N_cyt/C_lum orientation (Figure 8A). During subsequent topogenesis, this translocon complex will continue to partition the remaining regions of Nrf1 into the lumen or extraluminal environments around and within the ER membrane.
In addition to the above co-translational topogenesis, Nrf1 may be subject to a number of post-translational rearrangements even while the nascent CNC-bZIP protein remains engaged with the ribosome and translocon Sec61. Our glycosylation-mapping mutagenesis study and membrane proteinase protection reactions suggest that Nrf1 is reorganized within the ER membrane to possess intermediate topologies; they probably arise through post-translational modifications (e.g. glycosylation and deglycosylation) and other folding processes within the ER. These post-translational mechanisms enable the extra-TM1 regions in the NTD, SAS, CRACs and NHB2, to be either reinserted into membranes or repositioned from the lumen into the cytoplasm and/or nucleoplasm (Figures 8B and 8C). It should be noted that this topological reorganization is likely to be directed by the combined effects of protein–lipid interactions, the most important of which entails the amphipathic lipid membrane and several amphipathic amino acid α-helical regions within Nrf1. These amphipathic α-helical regions within Nrf1 may lie flat on the plane of membrane lipid bilayers, with its hydrophobic surface situated in the core hydrocarbon interior and its hydrophilic surface positioned on either the cytoplasm or in the luminal lipid–water interfaces. Furthermore, membrane lipid compositions may also modulate the topology of Nrf1 through ferrying it in the lipid plane between cholesterol-rich DRM and non-DRM membranes. The lipid composition of membranes may cause Nrf1 to flip-flop across the bilayers as the lipids move between the inner and outer leaflets. In addition, it is possible that the topology of Nrf1 is also modulated by internal interactions of TM1 with other amphipathic α-helical regions within the CNC-bZIP protein. Although this hypothesis requires confirmation, we predict that these interactions enable Nrf1 to form a local α-helical hairpin or bundle structure across membranes.

**Membrane topology of Nrf1 is determined by its NTD**

Nrf1 is an integral polytopic protein with a hydrophobic TM1 region, and several amphipathic α-helical regions that seem to be bis-span, mono-span or half-span the membrane lipid bilayer (Figures 8B and 8C). We have presented evidence suggesting that the membrane topology of Nrf1 is determined by its NTD. Within this domain, the TM1 α-helix represents a dominant membrane-interactive region with a hydrophobic surface and an amphipathic surface (see Supplementary Figure S1). The polar amino acids Thr8, Gln15, Glu16, Thr18 and Ser20 in TM1 may regulate its N$_{cyt}$/C$_{lum}$ orientation within the membrane and the extent to which it interacts with both the amphipathic membrane lipids and other amphipathic amino acid α-helix regions. The SAS α-helix (residues 31–50) is likely to be integrated into the membranes in co-operation with TM1, but its orientation may also be modulated by a putative Pro$_{37}$–Pro$_{38}$ kink within this region. We postulate that the CRACs subdomain (residues 62–82) interacts with cholesterol-rich DRMs, and thus the activity of Nrf1 may be influenced by the lipid content of membranes or it may even be involved in lipid sorting and membrane trafficking, as has been described for caveolin-1 [54]. Topological orientation of Nrf1 around the membrane is also controlled by its NHB2 region (residues 82–106) as it could serve as a dynamic switch anchoring the CNC-bZIP protein either on the membrane luminal leaflets or across membranes (Figure 8). In addition, the membrane topology of Nrf1 will also be modulated by other amphipathic subdomains (e.g. TMi and TMc) and by glycosylation of the NST domain [35]. It should be noted that membrane topology of Nrf1 is established within and around the ER before it is transported into the inner nuclear envelope membrane, because we observed significant differences in the time course of digestion of Nrf1 proteins in PK protection reactions with intact ER fractions, but not with entire nuclear membrane fractions (Supplementary Figure S5).

**The NTD of Nrf1 may confer a role in control of lipid homeostasis**

The ability of Nrf1 to induce ARE-driven gene expression is a consequence of its CNC-bZIP DNA-binding domain and its AD1 and AD2 transactivation domains [9,10,35]. Gene-knockout experiments have shown that Nrf1 is required for mouse development and growth, as well as the maintenance of liver function [19,20,22]. The markedly different phenotypes resulting from knockout of Nrf1 and Nrf2 indicate that the former transcription factor fulfills a unique role that the latter cannot perform. Among the many possible explanations for the distinct phenotypes, one possibility is that the NTD of Nrf1, which is not represented in Nrf2, endows Nrf1 with a unique function. In hepatocytes from liver-specific Nrf1$^{-/-}$ mice, a marked increase in the number of lipid vesicles and proliferation of smooth ER has been reported [15,22]. By contrast, the livers of Nrf2$^{-/-}$ mice show no phenotype under normal laboratory conditions [24]. It has been speculated that the abnormalities in the liver-specific Nrf1-null mice result from enhanced expression of the cytochrome P450 4A enzymes responsible for ω-oxidation of fatty acids [22]. We now hypothesize that the NTD of Nrf1 may be involved in controlling lipid homeostasis and membrane trafficking of the factor, because it shares conservation with the N-terminus of ArfGAP1, a Golgi-associated protein that contains an amphipathic ALPS (α-helix lipid-packing sensor) for membrane curvature [55]. Furthermore, like other membrane proteins [48,55–57], the amphipathic regions in Nrf1 could be organized to accommodate both polar and apolar microenvironments of membrane lipid bilayers. Once these amphipathic regions are integrated as TM α-helices, they might in turn cause a disordering of membrane lipid organization, primarily by acting as a simple barrier and restricting the diffusion of neighbouring lipids. Thus we propose that some membrane-interactive amphipathic regions within NTD and/or other domains of Nrf1 may enable the CNC-bZIP protein to play a role in membrane-dependent biological processes.

**AUTHOR CONTRIBUTION**

Yiguo Zhang designed and performed all experiments; Yiguo Zhang and John Hayes analysed experimental data and wrote the paper.

**ACKNOWLEDGEMENTS**

We thank Dr Changjiang Dong (Centre for Biomolecular Science, University of St Andrews, St Andrews, Scotland, U.K.) for his expert advice regarding the structural biology of membrane proteins. We thank Professor Ronald T. Hay (College of Life Sciences, University of Dundee, Dundee, Scotland, U.K.) for his help and interest in the present study. Y.Z. thanks the 111 Project for Biomechanics and Tissue Repair Engineering, Chongqing University, China, for travel support.

**FUNDING**

This work was supported by the Association for International Cancer Research (grant numbers 06-015, 09-0254).

**REFERENCES**


7 Nioi, P. and Hayes, J. D. (2004) Contribution of NAD(P)H:quinone oxidoreductase 1 to protection against carcinogenesis, and regulation of its gene by the Nrf2 basic-region leucine zipper and the arylhydrocarbon receptor basic helix-loop-helix transcription factors. Mutat. Res. 555, 149–171


SUPPLEMENTARY ONLINE DATA
Identification of topological determinants in the N-terminal domain of transcription factor Nrf1 that control its orientation in the endoplasmic reticulum membrane

Yiguo ZHANG*† and John D. HAYES*
*Biomedical Research Institute, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, Scotland, U.K., and †Laboratory of Cell Biochemistry and Gene Regulation, College of Bioengineering and Life Sciences, University of Chongqing, Chongqing 400044, China

Figure S1  Four wheels represent putative amphipathic α-helix structures within the NTD

Four distinct α-helix structures, formed by amino acids 7–24 (TM1), 31–50 (SAS), 55–80 (CRACs) and 89–106, that comprise a C-terminal 70% portion of NHB2, are predicted to exist in the NTD. In these wheels, the hydrophobic and hydrophilic surfaces are separated by a bold broken line. All the circled single letters represent amino acid residues with different physicochemical properties, of which all hydrophobic and apolar residues are represented as white letters on a black background, whereas acidic, basic and charged residues are placed on red, dark blue and light blue backgrounds respectively, and other polar residues sit on a white background. The net basic charged α-helix structure may be folded by amino acids covering the putative CRAC sequences (amino acids 55–80) and overlaps with a basic-rich N-terminal portion (amino acids 81–90) of NHB2. In the region, Leu62–X5–Tyr67–X4–Lys70 and Leu74–X2–Tyr77–X3–Arg81 represent two putative CRACs. A bold broken line separates the net hydrophobic and hydrophilic surfaces in this wheel.

1 To whom correspondence should be addressed (email y.z.zhang@dundee.ac.uk).
Figure S2  The intensity of the subcellular staining for wild-type Nrf1
Figure S3  The intensity of the subcellular staining for three Nrf1 mutants
Figure S4  SAS and CRACs are associated with membranes

(A) After COS-1 cells reached ~ 70% confluence in a 100-mm dish, they were transfected with Nrf1 expression constructs (6 μg of DNA for 6 h). Thereafter, the cells were changed into fresh complete medium for a further 24 h, before subcellular fractionations were performed. Equal 15 μg portions of protein were applied to each electrophoresis sample well from the cytosolic (C), ER, microsome-rich membrane (M) and the intact nuclei (N) fractions, as well as total lysates, and were resolved by SDS/NuPAGE Tris/Bis in a gel containing 4–12% polyacrylamide. Following electrophoresis, the samples were immunoblotted with antibodies against the V5 epitope, actin and CRT. (B) COS-1 cells expressing either ectopic Nrf1^{Δ2-30} or Nrf1^{Δ2-50} were subjected to sequential subcellular fractionation using centrifugation at different speeds of 1.3K, 5K, 8K, 17K and 100K (K indicates 1000 g). Aliquots of different pellet fractions (50 μg of protein) were incubated with 100 μg/ml (+) of PK in the presence (+) or absence (−) of 1% TX for 30 min on ice. The reaction products were resolved by a 4–12% SDS/NuPAGE Tris/Bis gel and visualized by Western blotting as described above. The molecular mass in kDa is indicated on the left-hand side of each gel.
Figure S5  Time-dependent membrane PK protection of Nrf1 in the ER and nuclear envelope

The upper two graphs show membrane PK protection of Nrf1 in intact ER (A) and nuclear envelope (B) over a 60 min period. In addition, COS-1 cells expressing Nrf1Δ31-80 (C) or Nrf1Δ81-106 (D) were subjected to subcellular fractionation, followed by intact ER and nuclear membrane PK protection reactions. For a detailed description of these data, see the legend of Figure 7 in the main text. For (C and D), the molecular mass in kDa is indicated on the left-hand side.
Figure S6 Membrane topology of Nrf1 is controlled by its NTD

(A) COS-1 cells were transfected with an expression construct for Nrf1, Nrf2 and Nrf3 as well as an empty pcDNA3.1/V5 His B vector. Approx. 24 h following transfection, the cells were transferred to complete DMEM medium containing 25 mmol/l or 0 mmol/l glucose for an additional 18 h. Subsequently, the cells were harvested and lysed. Proteins in samples were separated in a 4–12% SDS/NuPAGE Tris/Bis gel followed by immunoblotting with primary antibodies against either an Nrf1b polypeptide (upper panel) or a V5 epitope (middle and lower panels). The latter antibody-blotted nitrocellulose membrane was exposed to X-ray film for 5 s (middle panel) and 30 s (lower panel). (B) Samples for membrane proteinase protection reactions were prepared as described in the legend of Figure 7 in the main text. Aliquots of protein (50 μg in 40 μl of 1 X isotonic buffer) purified from the intact ER fractions of COS-1 cells expressing Nrf1, or its mutants, were incubated with PK at a final concentration of 50 μg/ml (right-hand panels) or 100 μg/ml (left-hand panels) for 30 or 60 min. These reaction products (5 μg of protein) were electrophoretically separated and visualized by Western blotting with either anti-Nrf1b antibodies (upper panel) or anti-V5 antibodies (middle panel). The former antibody-blotted nitrocellulose membrane was stripped and re-probed with antibodies against CRT (lower panel). The intensity of the Western blots was calculated using the Quantity One® software. The blot values of Nrf1 in the upper gel were divided by its corresponding sample CRT value in the lower panel, and then the resulting numbers were normalized to the net non-digested amounts and shown in the upper graphs. For the gels, the molecular mass in kDa is indicated on the left-hand side.