research papers

The structure of the Bach2 POZ-domain dimer reveals an intersubunit disulfide bond

Gareth O. Rosbrook,[a] Mark A. Stead,[a] Stephen B. Carr[b] and Stephanie C. Wright[a]*

[a]Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, England, and [b]Research Complex at Harwell, Rutherford Appleton Laboratory, Harwell Oxford, Didcot, Oxon OX11 0FA, England

Correspondence e-mail: s.c.wright@leeds.ac.uk

Bach2 is a transcriptional repressor that is expressed during specific stages of B-cell development and in neuronal cells. It plays a critical role in modulating class-switch recombination during the differentiation of mature B cells to antibody-secreting plasma cells and it is also an important regulator of apoptotic responses to oxidative stress. Bach2 has been implicated both as an oncogene and as a tumour suppressor in human malignancy. The interaction of Bach2 with its target genes is mediated *via* its basic leucine-zipper region, whereas the N-terminal POZ domain recruits transcriptional co-repressors and class II histone deacetylases. Here, the crystal structure of the human Bach2 POZ domain is reported at 2.1 Å resolution. The Bach2 POZ-domain dimer resembles the POZ-domain dimers of the POZ zinc finger transcription factors and dimerization is independent of an N-terminal region that has previously been implicated in the dimerization of the POZ basic leucine-zipper protein Bach1. The Bach2 POZ domain crystallized in two forms which differed by the presence of an intersubunit disulfide bond. The intersubunit disulfide bond is present both in bacterially expressed Bach2 POZ domain in solution and in protein expressed in trans-fected eukaryotic cells. These crystal structures will be relevant for understanding the regulation of Bach2 in response to oxidative stress and for the design of therapeutics that target the Bach2 POZ domain in human malignancy.

1. Introduction

The Bach (**b**ric-à-brac, tramtrack and broad complex **a**nd **c**ap'n'collar **h**omology) family transcription factors Bach1 and Bach2 are CNC (**c**ap'n'**c**ollar) type bZip (**b**asic leucine **zip**per) proteins that contain an N-terminal BTB [**b**ric-à-brac, **t**ramtrack and **b**road complex; also known as POZ (**po**xvirus and **z**inc finger)] domain (Oyake *et al.*, 1996) that mediates transcriptional repression. Bach1 is ubiquitously expressed, whereas Bach2 is found mainly in neuronal cells and during specific stages of B-cell differentiation. Bach1 and Bach2 interact with the small Maf (**m**usculo**a**poneurotic **f**ibrosarcoma) proteins *via* their leucine-zipper regions to form heterodimers that bind MARE (**Ma**f **r**ecognition **e**lement) sequences within target genes (Oyake *et al.*, 1996; Muto *et al.*, 1998).

Bach2 plays a role in the late stages of B-cell development and is also a critical component of the transcription-factor network that regulates the terminal differentiation of mature B cells to antibody-secreting plasma cells (reviewed in Igar-ashi *et al.*, 2007). Bach2 is highly expressed throughout the pro-B to mature B-cell stages (Muto *et al.*, 1998), during which

it represses transcription of the *Blimp-1* (**B** **l**ymphocyte-**i**nduced **m**aturation **p**rotein-**1**) gene; Blimp-1 is required for terminal differentiation to plasma cells. Bach2 levels decrease upon B-cell activation following antigen exposure, leading to *Blimp-1* induction and the concomitant differentiation of B cells into IgM antibody-producing plasma cells. During antibody responses, class-switch recombination (CSR) of the immunoglobulin genes leads to the production of additional antibody classes, such as IgG, which have different effector functions. CSR is mediated by the transient induction of activation-induced cytidine deaminase (AID) following B-cell activation; importantly, AID is no longer expressed in terminally differentiated plasma cells. The reduction in Bach2 levels is heterogenous between individual activated B cells and leads to a variation in the timing of *Blimp1* induction and in the rate of differentiation to plasma cells; this allows a time window that delays differentiation in those cells with higher Bach2, thereby enabling AID-mediated class-switch recombination to occur (Muto *et al.*, 2010). Mice that are deficient in Bach2 have reduced numbers of mature B cells, enhanced rates of differentiation to plasma cells and decreased CSR (Muto *et al.*, 2004). Bach2 thus plays a crucial role in repressing *Blimp1* expression in activated B cells long enough to delay differentiation and thereby allowing CSR to occur, and the heterogenous decline in Bach2 explains the stochastic nature of plasma-cell differentiation and of CSR.

The activity of Bach2 is modulated by its nucleocytoplasmic shuttling, which is regulated by the PI3 (**p**hospho**i**nositide **3**) kinase pathway and by oxidative stress. Following B-cell activation, the phosphorylation of Ser521 by PI3 kinase leads to the movement of Bach2 from the nucleus to the cytoplasm, thereby promoting plasma-cell differentiation and inhibiting CSR (Yoshida *et al.*, 2007). Bach2 also plays an important role in modulating the cellular responses to oxidative stress, in part by repressing transcription of the *haem oxygenase-1* (HO-1) gene (Yoshida *et al.*, 2007), which has anti-apoptotic functions and cytoprotective effects against oxidative injury. Low levels of oxidative stress lead to the phosphorylation of Bach2 and retention in the cytoplasm, resulting in derepression of the anti-apoptotic *HO-1* gene and allowing cells to repair and survive. In contrast, high levels of oxidative stress lead to the dephosphorylation and nuclear import of Bach2, resulting in repression of *HO-1* and concomitant apoptosis. Bach2 is thought to regulate apoptotic responses to the oxidative stress that is associated with B-cell development and also to modulate responses to chemotherapeutic agents that induce oxidative stress (Kamio *et al.*, 2003). The cytoplasmic retention of Bach2 is directed by its C-terminal cytoplasmic localization signal (Hoshino *et al.*, 2000), whereas the N-terminal POZ domain is required for the association of Bach2 with PML (**pro**myelocytic **l**eukaemia protein) bodies in the nucleus (Tashiro *et al.*, 2004).

Bach2 has been implicated both as an oncogene and as a tumour suppressor in human malignancies. The BCR-ABL (**b**reakpoint **c**luster **r**egion-**Ab**el**s**on) fusion protein that is expressed in chronic myelogenous leukaemia activates signalling pathways that lead to the increased phosphorylation and nuclear export of Bach2, resulting in enhanced cell survival owing to resistance to oxidative stress (Yoshida *et al.*, 2007). Treatment of leukaemic cells with the BCR-ABL inhibitor ST1571 leads to elevated nuclear Bach2 and increased apoptosis; apoptosis is further enhanced by simultaneous treatment with oxidative stressors. Expression of the anti-apoptotic *HO-1* gene is high in some malignancies (Hara *et al.*, 1996) and the resistance of myeloma cells to ROS-induced (**r**eactive **o**xidant **s**pecies) apoptosis is associated with increased HO-1 levels and reduced Bach2 levels (Zhou *et al.*, 2005). In contrast to its role as a tumour suppressor, Bach2 has been implicated as an oncogene in leukaemic cells that have chromosome translocations that juxtapose the *Bach2* gene to the IgH locus (Kobayashi *et al.*, 2011). Additionally, Bach2 overexpression has been reported in some non-Hodgkin's lymphoma cell lines that carry an Ig-Myc (**my**elo**c**ytoma) translocation (Kobayashi *et al.*, 2011) and overexpression of Bach2 has been correlated with a poor prognosis in epithelial ovarian cancer (Motamed-Khorasani *et al.*, 2007). The contrasting roles of Bach2 in different human malignancies is most likely to reflect its opposing functions in preventing differentiation and promoting apoptosis.

The POZ domain is a protein–protein interaction domain that is found in approximately 40 mammalian transcription factors (POZ-TFs); POZ domains mediate interactions with transcriptional co-regulators and also direct specific interactions of the POZ-TFs with each other (reviewed in Stogios *et al.*, 2005). Most POZ-TFs contain a zinc-finger DNA-binding domain and several members of this family have been extensively studied owing to their role in human malignancy. For example, the BCL6 (**B**-**c**ell **l**ymphoma **6**) oncogene is overexpressed in follicular lymphoma and in diffuse large-cell B-cell lymphoma and structural studies of the interactions of the BCL6 POZ domain with its transcriptional co-repressors B-CoR (**B**CL6-interacting **co**-**r**epressor; Ghetu *et al.*, 2008) and SMRT (**s**ilencing **m**ediator for **r**etinoid and **t**hyroid hormone receptors; Ahmad *et al.*, 2003) led to the design of inhibitors that may be developed into therapeutics (Cerchietti *et al.*, 2010, 2009). The target gene repertoire of the POZ-TFs is increased *via* their heteromeric interactions with each other; for example, the interaction between BCL6 and the POZ-TF Miz1 (**M**yc-**i**nteracting **z**inc finger **1**) leads to the recruitment of BCL6 to Miz1 target genes (Phan *et al.*, 2005) and is important in suppressing Miz1 targets during the germinal centre reaction in B cells. The Bach2 POZ domain interacts with the POZ zinc-finger protein MAZR (**M**yc-**a**ssociated **z**inc-finger protein **r**elated factor; Kobayashi *et al.*, 2000) and Bach2 also interacts with BCL6 in a POZ-independent manner (Ochiai *et al.*, 2008). The Bach2 POZ domain represses transcription *via* the recruitment of class II histone deacetylases and SMRT (Hoshino *et al.*, 2007), although the POZ domains of the POZ bZip proteins are not well characterized and detailed structures of their interactions are not available.

We report the crystal structure of the human Bach2 POZ domain to 2.1 Å resolution. The overall structure resembles the POZ domains of the POZ zinc-finger factors and, in contrast to a previous report (Ito *et al.*, 2009), suggests that

dimerization of the POZ bZip POZ domains does not require an N-terminal hook. The structure of the Bach2 POZ domain reveals an intersubunit disulfide bond that is also observed in purified protein in solution and *in vivo*.

2. Materials and methods

2.1. Cloning

DNA fragments encoding the human Bach2 POZ domain (Bach2 residues 9–129 and residues 2–132) were amplified by PCR from a human placental cDNA library and inserted into the bacterial expression vector pGEX-6P-1 (GE Healthcare) as *Bam*HI–*Xho*I restriction-enzyme fragments. The resulting plasmids encoded fusion proteins containing an N-terminal GST tag and a PreScission protease cleavage site. The human Bach2 POZ domain C20S mutation was generated by PCR with Phusion High-Fidelity DNA polymerase (Finnzymes) using the oligonucleotides GAGTCCACAGTCCACTCCACCAACATCCTCCTGGGC and ATACACATACATGGGGGAGTCAGGCTTCTCATCCAC. The C125S mutation was generated using the oligonucleotides CTGGAGGACTCCTCCTTCAGCTTCCTGCAGACC and GTTGTGCATGCGCAGGAACTCAGCACAGCG.

The mouse Bach2 POZ domain (residues 2–132) was amplified by PCR from a mouse ES cell cDNA library and was cloned into the mammalian expression vector pcDNA3 (Invitrogen); an oligonucleotide encoding a FLAG-tag was inserted at the N-terminus of the POZ domain and an oligonucleotide encoding the c-Myc nuclear localization signal was inserted at the C-terminus. The mouse Bach2 POZ domain C20S mutation was generated as above using the oligonucleotides GAGTCCACAGTCCACTCTGCCAACATCCTCCTG and ATATACATACATGGGGGAGCCAGGCTTCTCTC.

2.2. Protein expression and purification

The plasmids were transformed into *Escherichia coli* BL21 (DE3) pLysS cells using the heat-shock method and plated on Luria–Bertani agar plates containing 34 mg l^{-1} chloramphenicol and 100 mg l^{-1} ampicillin. Cells were cultured overnight with shaking at 310 K in 2TY medium containing the same antibiotics. 100 ml of the overnight culture was subsequently inoculated into 10 l 2TY medium supplemented with 100 mg l^{-1} ampicillin and the cells were grown at 310 K with shaking to an optical density at 600 nm of 0.6–0.8. Expression of recombinant protein was induced with isopropyl β-D-1-thiogalactopyranoside at a final concentration of 0.1 mM and the cells were incubated at 289 K for 16 h. The cells were harvested by centrifugation and the cell pellets were stored at 193 K until required.

A bacterial pellet from a 10 l culture was resuspended in lysis buffer [PBS containing 0.1% Triton X-100 and 5 mM dithiothreitol (DTT) pH 7.5] and the cells were lysed using a French press. The lysate was clarified by centrifugation at 39 000g for 20 min and the supernatant was passed through a 0.45 µm filter. Fusion proteins were bound to glutathione-Sepharose 4B that had been pre-equilibrated with lysis buffer.

Nonspecifically bound proteins were removed from the glutathione-Sepharose 4B by alternating washes in lysis buffer and in lysis buffer supplemented with 1 M NaCl. The glutathione-Sepharose 4B was further rinsed with 20 mM Tris–HCl, 75 mM NaCl, 5 mM DTT pH 7.5 and the GST tag was removed by cleavage with PreScission protease. Bach2 POZ domains were further purified by size-exclusion chromatography on a pre-equilibrated HiLoad Superdex 75 26/60 column (GE Healthcare) in 20 mM Tris–HCl, 150 mM NaCl, 5 mM DTT, 5% glycerol pH 7.9. Purified Bach2 POZ domains eluted at a concentration of 0.2 mg ml^{-1} and were concentrated to 12 mg ml^{-1} using Amicon Ultra centrifugal concentrators (Millipore).

2.3. Protein crystallization

Crystallization trials were set up in 96-well Corning 3552 crystallization trays using an Oryx 6 robot (Douglas Instruments Ltd). A variety of commercially available sparse-matrix crystallization screens were used and crystals of the human Bach2 POZ domain (residues 9–129) were obtained in 0.2 M ammonium sulfate, 0.1 M bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane pH 5.5, 25%(w/v) PEG 3350 at 291 K using a protein concentration of 6.5 mg ml^{-1}. Optimization of the crystallization conditions was carried out by sitting-drop vapour diffusion in MVD-24 plates (NBS Biologicals). Crystals with average dimensions of $100 \times 30 \times 30$ µm were grown by mixing 2 µl protein solution (9 mg ml^{-1}) with 2 µl reservoir solution [0.2 M ammonium sulfate, 0.1 M 2-(N-morpholino)-ethanesulfonic acid pH 5.5, 4%(w/v) PEG 3350] and incubating at 291 K for 4–5 d. Crystals were harvested with a nylon cryoloop and transferred into a 5 µl drop of 0.2 M ammonium sulfate, 0.1 M 2-(N-morpholino)ethanesulfonic acid pH 5.5, 4%(w/v) PEG 3350, 17.5% glycerol. They were soaked for 30 s in this solution before being cooled to 100 K.

2.4. Data collection, structure determination and refinement

X-ray diffraction data were collected on beamline I02 at Diamond Light Source, processed using *MOSFLM* (Leslie, 1992) and scaled using *SCALA* (Evans, 2006). The human Bach2 POZ domain (residues 9–129) crystallized in space group $P2_1$. A Matthews coefficient corresponding to a bulk-solvent content of 52% suggested the presence of six molecules in the asymmetric unit. The structure was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007) with the human Bach1 POZ domain (PDB entry 2ihc; Structural Genomics Consortium, unpublished work) as the search model.

Initial models were built using *ARP/wARP* (Cohen *et al.*, 2008) and underwent density modification using *Parrot* (Zhang *et al.*, 1997). Several rounds of iterative model building and refinement were performed using *Coot* (Emsley *et al.*, 2010) and *REFMAC*5 (Murshudov *et al.*, 2011); 5% of the data were excluded from refinement and used for cross-validation. Later rounds of refinement included TLS refinement with two groups per POZ-domain monomer (Painter & Merritt, 2006); TLS parameters were calculated using the *TLSMD* web server

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	Bach2 form I (Cys20 reduced)	Bach2 form II (Cys20 oxidized)
Crystal parameters		
Space group	$P2_1$	$P2_1$
Unit-cell parameters (Å, °)	$a = 55.60$, $b = 96.97$, $c = 82.63$, $\alpha = 90.00$, $\beta = 101.06$, $\gamma = 90.00$	$a = 55.68$, $b = 96.95$, $c = 82.48$, $\alpha = 90.00$, $\beta = 101.06$, $\gamma = 90.00$
Data collection		
Resolution (Å)	54.57–2.10 (2.21–2.10)	48.48–2.20 (2.32–2.20)
Wavelength (Å)	0.9795	0.9795
R_{merge}† (%)	9.3 (72.6)	10.5 (71.8)
$R_{\mathrm{p.i.m.}}$‡ (%)	6.1 (47.7)	6.9 (47.7)
$\langle I/\sigma(I)\rangle$	9.2 (2.9)	8.9 (1.7)
No. of unique reflections	49842 (7295)	43314 (6332)
Multiplicity	3.1 (3.1)	3.1 (3.1)
Completeness (%)	99.2 (99.7)	99.2 (99.7)
Refinement		
Resolution (Å)	40.55–2.10	48.48–2.20
R§ (%)	21.0	22.3
R_{free}¶ (%)	25.7	27.3
R.m.s.d. stereochemistry††		
Bond lengths (Å)	0.011	0.011
Bond angles (°)	1.3	1.3
No. of protein atoms	5710	5731
No. of water molecules	276	216
Average B factor (Å^2)	38.6	31.7
Ramachandran analysis‡‡ (%)		
Favoured	97.6	96.1
Allowed	2.4	3.9
Disallowed	0.0	0.0
PDB code	3ohu	3ohv

† $R_{\mathrm{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the integrated intensity of a given reflection and $\langle I(hkl)\rangle$ is the mean intensity of multiple corresponding symmetry-related reflections. ‡ $R_{\mathrm{p.i.m.}} = \sum_{hkl}[1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the integrated intensity of a given reflection, $\langle I(hkl)\rangle$ is the mean intensity of multiple corresponding symmetry-related reflections and N is the multiplicity of a given reflection. § $R = \sum_{hkl} ||F_{\mathrm{obs}}| - |F_{\mathrm{calc}}|| / \sum_{hkl} |F_{\mathrm{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. ¶ R_{free} is R calculated using a random 5% of data excluded from the refinement. †† R.m.s.d stereochemistry is the deviation from ideal values. ‡‡ Ramachandran analysis was carried out using *MolProbity* (Chen *et al.*, 2010).

(http://skuld.bmsc.washington.edu/~tlsmd). Geometry was analysed using *PROCHECK* (Laskowski *et al.*, 1993) and *MolProbity* (Chen *et al.*, 2010). R.m.s.d values for the POZ domains were calculated using the *SuperPose* (Maiti *et al.*, 2004) server (http://wishart.biology.ualberta.ca/SuperPose).

OMIT maps were generated using *SFCHECK* (Vaguine *et al.*, 1999) and images of the protein structures were prepared using *CCP4MG* (Potterton *et al.*, 2004).

2.5. SDS–PAGE under reducing and nonreducing conditions

Wild-type, C20S, C125S and C20S/C125S mutant human Bach2 POZ domains were expressed in *E. coli* and purified as described above. Samples for reduction were treated with β-mercaptoethanol to a final concentration of 5% and samples for oxidation were treated with hydrogen peroxide to a final concentration of 0.05–5 mM for 1 h at room temperature. All samples were heated for 5 min at 368 K in 62.5 mM Tris–HCl, 2% SDS, 10% glycerol, 0.05% Bromophenol Blue and analysed by SDS–PAGE (Laemmli, 1970) using 15% gels. Proteins were visualized by staining with Coomassie R-250.

2.6. Transfection of eukaryotic cells

COS-7 cells were transfected with expression vectors encoding FLAG-tagged mouse Bach2 POZ domains (wild type and C20S mutant) using the calcium phosphate precipitation method (Wigler *et al.*, 1979). Cells were harvested 48 h after transfection and lysed in 50 mM Tris–HCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 pH 7.4. Samples were electrophoresed on SDS–polyacrylamide gels and transferred to a nitrocellulose membrane. The membrane was probed with 1:2000 rabbit polyclonal anti-FLAG antibody (Sigma–Aldrich) followed by 1:10 000 HRP-conjugated anti-rabbit IgG (Pierce) before being exposed to X-ray film.

3. Results

3.1. Structural features of the human Bach2 POZ domain

The Bach2 POZ domain (residues 9–129) crystallized in space group $P2_1$ with three POZ-domain dimers in the asymmetric unit. The overall structure of the Bach2 POZ dimer resembles the reported POZ dimers of the POZ zinc-finger transcription factors.

It was apparent upon solving the structure that the Bach2 POZ domain crystallized in two forms, form I and form II, which differ in the oxidation state of the Cys20 residue located at the N-terminus of the α1 helix in the dimerization interface (Table 1 and Fig. 1a). Both crystal forms grew under identical conditions. Cys20 is in a reduced state in all six Bach2 chains in the asymmetric unit of the form I crystals, whereas all Bach2 dimers in the form II crystals contain a Cys20 intersubunit disulfide bond. OMIT maps were calculated and confirmed the electron density in the vicinity of Cys20 (Supplementary Fig. S1[1]). A disulfide bond is also present between the Cys125 residues of two neighbouring POZ dimers in the form II crystal structure; the orientation of the other molecules in the asymmetric unit of this crystal is not compatible with the formation of Cys125 disulfide bonds. Cys125 is reduced in all chains of the form I crystals and all other cysteine residues are reduced in both forms I and II. The overall structures of the POZ-domain dimers in the form I and II crystals are otherwise identical, with an r.m.s.d. value of C$^\alpha$-atom positions of 0.2 Å, and there are no differences in the positions of the side chains in the vicinity of Cys20. We also crystallized a slightly longer version of the human Bach2 POZ domain (residues 2–132) using different crystallization conditions; although these crystals diffracted to only 3 Å resolution it was clear that this longer protein also crystallized in two forms that differed in the oxidation state of Cys20 (data not shown).

The overall organization of the Bach2 POZ domain is similar to the previously reported POZ-domain dimers of PLZF (**p**romyelocytic **l**eukaemia **z**inc **f**inger; Ahmad *et al.*, 1998), BCL6 (Ahmad *et al.*, 2003; Stead *et al.*, 2008), LRF (**l**eukaemia/lymphoma **r**elated **f**actor; Schubot *et al.*, 2006), Miz1 (Stead *et al.*, 2007; Stogios *et al.*, 2010) and Nac1 (**n**uclear
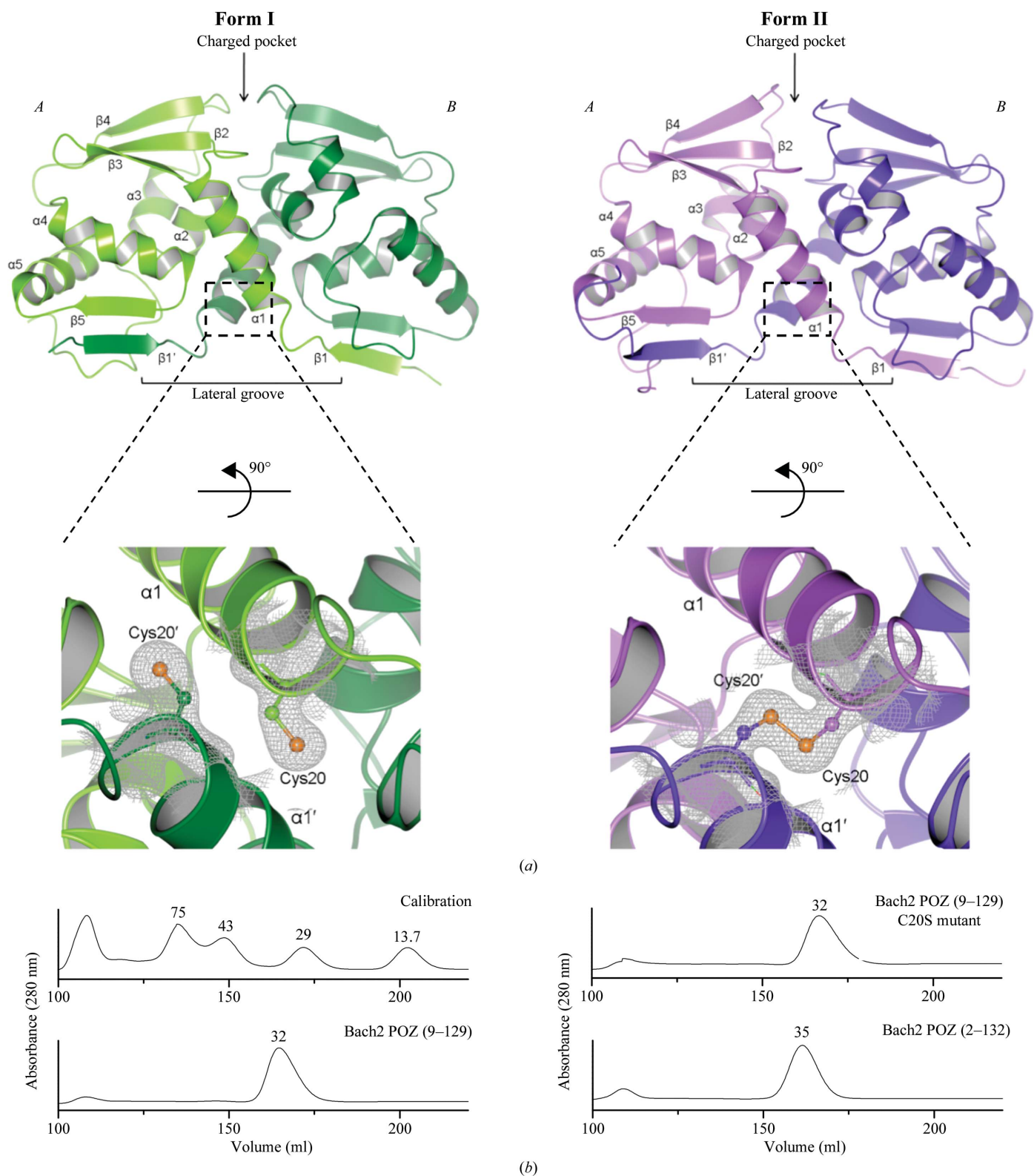
Figure 1
Structure of the human Bach2 POZ domain. (*a*) Ribbon representation of the form I (Cys20 reduced) and form II (Cys20 oxidized) structures of the human Bach2 POZ domain. Secondary-structure elements of the *A* subunit are labelled; a prime (′) indicates a secondary-structure element of the *B* chain. The lower images depict an enlargement of the region surrounding the Cys20 residue; S atoms are coloured orange. A $2F_o - F_c$ electron-density map is shown contoured at 1.5σ. The charged pocket and lateral grooves are indicated; these regions have previously been described in other POZ-domain proteins (Ahmad *et al.*, 1998, 2003). (*b*) Size-exclusion chromatography of wild-type and mutant Bach2 POZ domains using Superdex 75. The top left trace represents the elution profile of marker proteins (low-molecular-weight gel-filtration calibration kit; GE Healthcare). Sizes are indicated in kDa. The wild-type and mutant Bach2 POZ domains eluted at volumes corresponding to dimeric species.

*ac*cumbens associated **1**; Stead *et al.*, 2009), in which a central core of α-helices is flanked by short β-sheets. Compared with the Bach2 POZ domain (form I), the r.m.s.d. values for C$^\alpha$-atom positions are 2.6 Å for PLZF (PDB entry 1buo), 1.8 Å for BCL6 (PDB entry 1r28), 2.9 Å for LRF (PDB entry 2nn2) and 1.1 Å for Bach1 (PDB entry 2ihc). In agreement with reported POZ-domain structures, the Bach2 POZ domain is a domain-swapped dimer with two types of interaction interface: a central hydrophobic core that involves interactions from α1, α2 and α3 and that buries approximately 1200 Å^2 of surface area per monomer, together with two β-strand interfaces that comprise interactions between the β1 strand of one chain and β5 of the other and which bury an additional 1000 Å^2 of surface area per monomer (Fig. 1a).

The C-termini of the BCL6, LRF, PLZF and Miz1 POZ domains form an α-helix (α6) that lies adjacent to α5; however, as observed in Nac1 (Stead *et al.*, 2009) this region is unstructured in the Bach2 POZ domain. α6 is not part of the POZ-domain core fold and the residues within this region are the least conserved in both sequence and structure among the various POZ domains.

3.2. Dimerization of the Bach2 POZ domain does not require an N-hook

The POZ-domain sequences of the bZip transcription factors Bach1 and Bach2 are more closely related to each other than they are to those of POZ zinc-finger transcription factors (Fig. 2). The crystal structures of the human and mouse (Ito *et al.*, 2009) dimeric Bach1 POZ domains have recently been reported (PDB entries 2ihc and 2z8h, respectively; Structural Genomics Consortium, unpublished work; Ito *et al.*,

2009). In addition to the classic POZ-domain dimerization interfaces, the mouse Bach1 POZ-domain structure revealed an interaction between the N-terminal six residues of one chain and the very C-terminal residues of the opposite strand; these N-terminal residues that precede β1 have been dubbed the 'N-hook'. Analytical ultracentrifugation experiments showed that the wild-type Bach1 POZ domain is dimeric, whereas the sedimentation profile of a truncated protein that lacked the N-hook was interpreted as representing a monomeric species (Ito *et al.*, 2009). The N-terminal region of Bach2 (residues 1–9) contains an insertion of three residues compared with Bach1 (Fig. 2) and in the same analytical ultracentrifugation studies it was suggested that the sedimentation profile of purified Bach2 POZ domain represents that of a monomeric species. It was hypothesized that the N-hook is essential for dimerization of the Bach1 POZ domain and that Bach2 fails to dimerize because its N-terminus is unable to form this structure. It was therefore proposed that the mechanism of POZ-domain dimerization in the bZip factors differs from that observed in the POZ zinc-finger proteins.

In contrast to the reported study (Ito *et al.*, 2009), the human Bach2 POZ proteins (residues 2–132 and 9–129) that we expressed and purified both eluted as dimers during size-exclusion chromatography (Fig. 1b) and both crystallized as dimers using a variety of crystallization conditions (Fig. 1a and data not shown). This suggests that the N-terminal residues of Bach2 (residues 1–8) do not influence its dimerization, which is mediated solely by the classic POZ-domain interaction interfaces as observed in the POZ zinc-finger transcription factors. Of note, the crystal structure of the human Bach1 POZ domain (PDB entry 2ihc, residues 7–128) lacks the putative
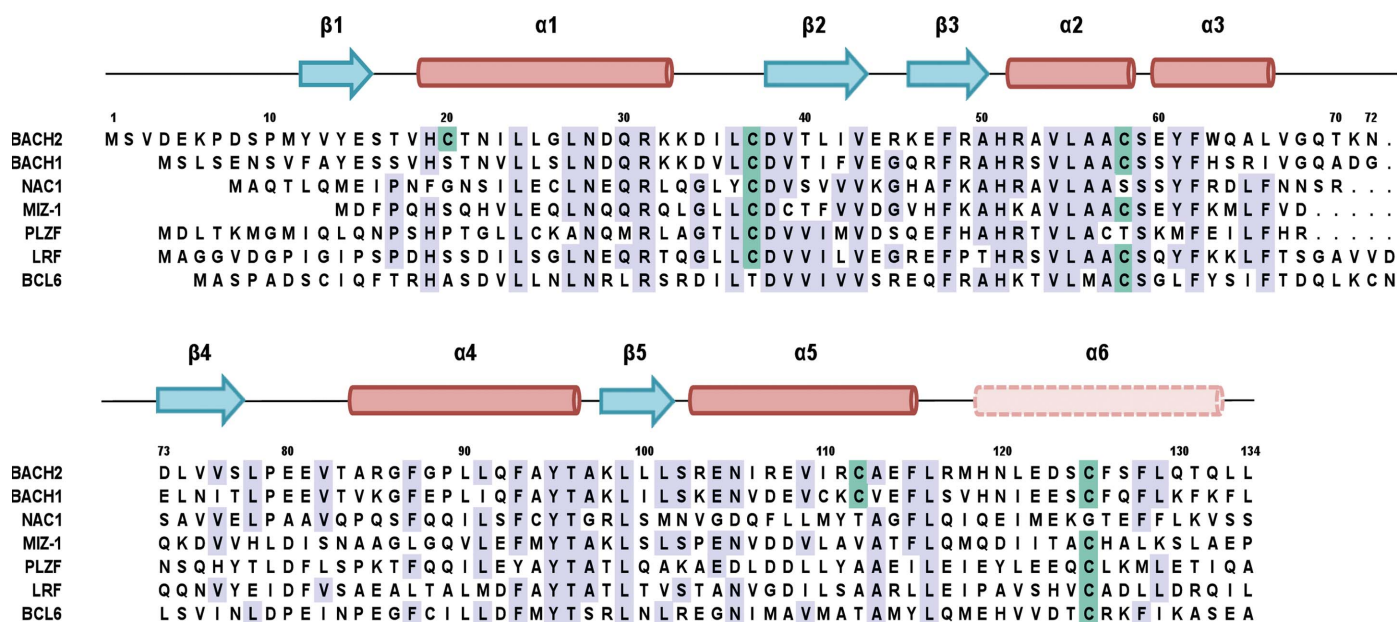
Figure 2
POZ-domain sequence alignment. Sequence alignment of a selection of human POZ-TF POZ domains; sequences were aligned using *ClustalW* (Thompson *et al.*, 1994). The observed secondary-structure elements of the Bach2 POZ-domain chains are indicated; α-helices are depicted as red cylinders and β-strands as blue arrows. The pink cylinder represents the sequence that corresponds to α6 in many POZ-TFs. All cysteine residues of the Bach2 POZ domain are highlighted in green and conserved cysteines of other POZ proteins are also indicated.

N-hook, thereby questioning the role of this region in regulating POZ-domain dimerization in the POZ bZip factors.

3.3. Formation of an intersubunit disulfide bond in the Bach2 POZ-domain dimer *in vitro* and *in vivo*

We determined whether the intersubunit disulfide bond observed in the human Bach2 POZ-domain crystals is also present in purified protein in solution and in protein expressed in eukaryotic cells *in vivo*. In order to abolish Cys20 disulfide bonds, we generated a human Bach2 POZ C20S mutant and purified recombinant protein that had been expressed in *E. coli*; the Bach2 POZ C20S mutant eluted solely as a dimer during size-exclusion chromatography (Fig. 1b). Purified wild-type and C20S Bach2 POZ domains were analysed by SDS–PAGE under both reducing and nonreducing conditions. Samples were either reduced by treatment with 20 mM DTT or oxidized by treatment with hydrogen peroxide. The wild-type POZ domain formed both dimers and higher-order oligomers under nonreducing conditions, with higher-order oligomerization being increased at higher concentrations of hydrogen peroxide (Fig. 3a). Dimerization under nonreducing conditions was significantly decreased in the C20S mutant

and higher-order oligomers were virtually absent (Fig. 3b). This suggested that Cys20 residues form intersubunit disulfide bonds within the POZ-domain dimer in solution and that an additional cysteine residue is most likely to be involved in the formation of disulfide bonds that link POZ-domain dimers into higher-order oligomers. Analysis of the crystal structure revealed that Cys125 is the only surface-exposed cysteine residue; we therefore made a C125S mutant and a C20S/C125S double mutant of the human Bach2 POZ domain. The C125S Bach2 POZ domain migrated as a dimer under oxidizing conditions, whereas all oligomerization was abolished in experiments with the C20S/C125S double mutant (Fig. 3c). The higher-order oligomerization of the wild-type human Bach2 POZ domain observed under nonreducing conditions in solution is therefore consistent with the formation of Cys20 disulfide bonds between the subunits of the POZ-domain dimer, together with Cys125 disulfide bonds that link the POZ-domain dimers into oligomers.

To determine whether the Bach2 POZ domain contains an intersubunit disulfide bond *in vivo*, we expressed FLAG-tagged wild-type and C20S mutant mouse Bach2 POZ domains in COS-7 cells and analysed cell lysates by Western blotting following SDS–PAGE under nonreducing and
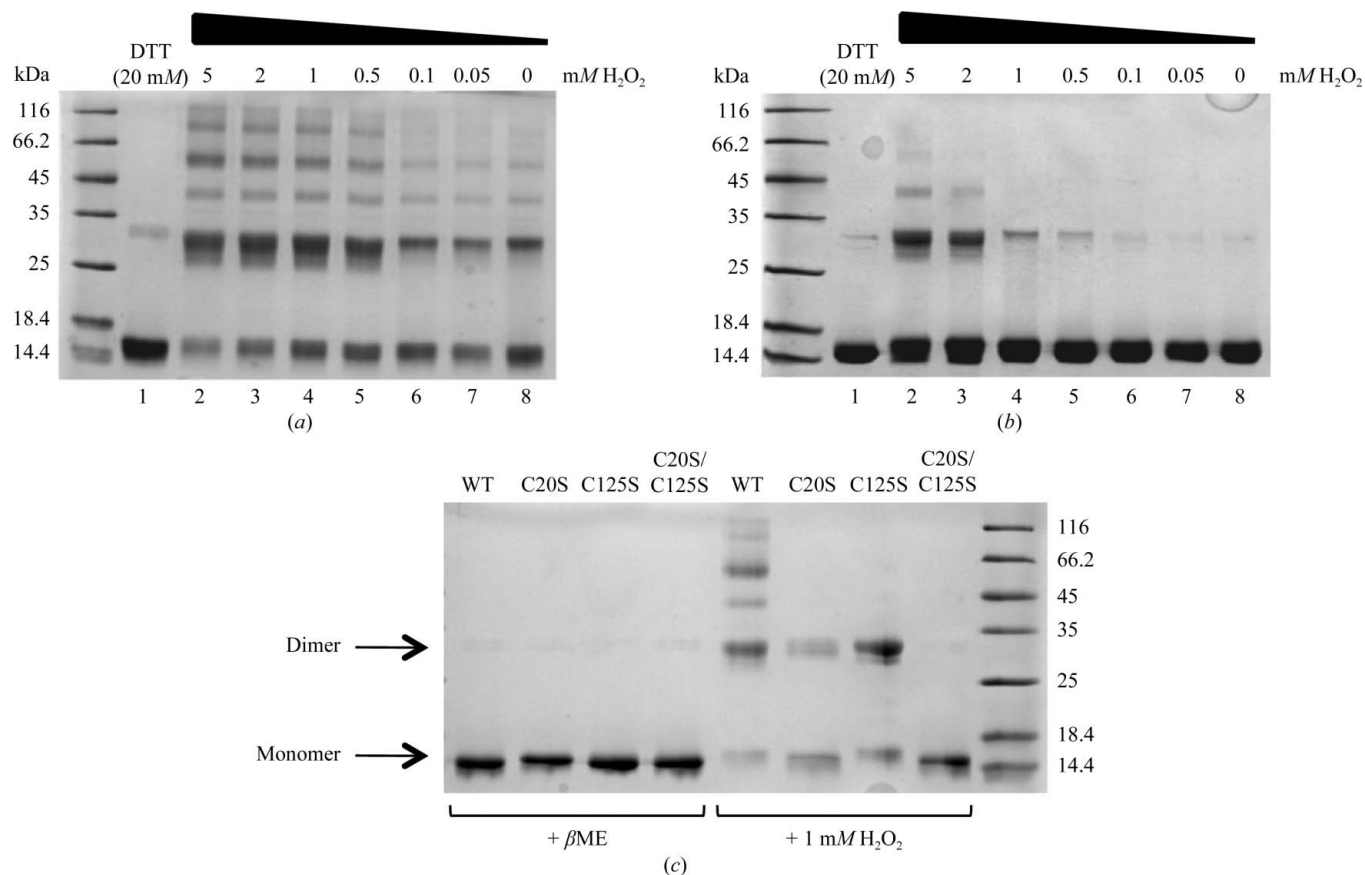
Figure 3
Disulfide-bond formation in the purified human Bach2 POZ domain in solution. (a) Analysis of purified wild-type Bach2 POZ-domain protein by SDS–PAGE. The sample in lane 1 contained 20 mM DTT and the samples in lanes 2–8 were treated with various concentrations of hydrogen peroxide for 1 h at room temperature and electrophoresed under nonreducing conditions. (b) Analysis of purified C20S Bach2 POZ-domain protein by SDS–PAGE. Samples were treated as described in (a). (c) Purified wild-type Bach2 POZ-domain protein, together with the C20S, C125S and C20S/C125S mutants, was analysed by SDS–PAGE. Reduced samples were treated with 5% β-mercaptoethanol (βMe). Oxidized samples were incubated with 1 mM hydrogen peroxide for 1 h at room temperature.

reducing conditions. Under nonreducing conditions the wild-type mouse Bach2 POZ domain migrated as both a monomer and a dimer, whereas the C20S mutant was monomeric (Fig. 4). This suggests that Cys20 mediates the formation of inter-subunit disulfide bonds in the Bach2 POZ domain *in vivo* and that the Cys125 disulfide bonds that we observed in bacterially expressed protein in solution were most likely to be a reflection of the high protein concentration *in vitro*.

4. Discussion

Reversible disulfide-bond formation is often used as a mechanism to regulate the activity of proteins involved in oxidative stress responses. Regulatory disulfide bonds are sometimes formed between distantly related cysteine residues, resulting in large conformation changes that modulate specific properties of the protein. For example, oxidation of the prokaryotic transcriptional regulator OhrR (**o**rganic **h**ydroperoxide **r**esistance **R**) by organic hydroperoxides leads to the formation of an intersubunit disulfide bond that results in a large conformation change; this weakens the interaction of OhrR with the DNA backbone and allows the derepression of genes that reduce the hydroperoxides to less toxic alcohols (Newberry *et al.*, 2007). Disulfide-bond formation may also directly stabilize subunit interactions within a protein or may influence protein–protein interactions. A disulfide bond is the main feature that stabilizes the dimerization interface of the activation domain of the bovine papilloma virus E2 protein (Sanders *et al.*, 2007); tight dimerization of this region abrogates its interaction with the viral E1 protein, thereby regulating E2 function in response to oxidative stress. Oxidative stress also leads to the inhibition of specific protein–protein interactions in the yeast transcriptional activator Yap1 (**Y**amaguchi sarcoma viral oncogene-**a**ssociated **p**rotein **1**; Wood *et al.*, 2004); disulfide-bond formation masks the nuclear export signal and prevents interaction with the CRM1 exporter, resulting in nuclear retention of Yap1 and the transcriptional activation of oxidative-stress response genes. Disulfide bonds may also be involved in directly mediating specific protein–protein interactions in response to oxidative stress: oxidation of the yeast tyrosine phosphatase Sdp1 (**s**tress-inducible **d**ual-specificity MAP kinase **p**hosphatase **1**) leads to the formation of a disulfide bond that forms part of a deep cleft involved in the specific recognition of substrate phosphotyrosine residues (Fox *et al.*, 2007). The intersubunit disulfide bond of the Bach2 POZ domain is not required for stabilization of the dimerization interface and does not lead to conformation changes within the POZ-domain dimer. The disulfide bond formed readily in crystals and in solution, even though the protein was purified in the presence of 5 m*M* DTT, and was observed in crystals under conditions in which other cysteines were reduced. Myoneurin, a protein expressed in muscle, is the only other POZ-TF with a cysteine residue located at the start of the α1 helix at a position corresponding to Cys20 in Bach2; the crystal structure of the myoneurin POZ-domain dimer (PDB entry 2vpk; C. D. O. Cooper, J. W. Murray, A. Bullock, A. C. W. Pike, F. Vondelft, P. Filippako-

poulos, E. Salah, A. Edwards, C. H. Arrowsmith, C. Bountra, J. Weigelt & S. Knapp, unpublished work) revealed an inter-subunit disulfide bond at this position. Other cysteine residues of the Bach2 POZ domain are conserved among most POZ-TF POZ domains, with the exception of Cys112, which is only found in the POZ bZip protein Bach1, a protein that is also regulated by oxidative stress.

The dimeric organization of the Bach2 POZ domain implies that this interaction will direct the higher-order association of Bach2–Maf heterodimers, analogous to the reported oligomerization of Bach1–Maf complexes (Igarashi *et al.*, 1998). Such interactions are predicted to facilitate the cooperative binding to multiple MARE sequences within target genes.

The physiological role of the intersubunit disulfide bond in the Bach2 POZ-domain dimer is presently unknown and it is not known whether it is part of the mechanism for sensing oxidative stress during B-cell development. It will be relevant to determine whether this disulfide bond affects the inter-actions of Bach2 with other POZ partners or with transcriptional co-repressors and whether it is involved in the mechanism that determines the subnuclear localization of Bach2 in response to oxidative stress.

The structure of the Bach2 POZ domain will aid in the development of therapeutics for those human malignancies in which Bach2 is overexpressed. In a manner analogous to the targeting of BCL6 in diffuse large B-cell lymphoma (Cerchietti *et al.*, 2009, 2010), peptides or small molecules could potentially be used to inhibit the interactions of the Bach2 POZ domain with histone deacetylases and the transcriptional co-repressor SMRT. Such approaches would require the characterization of the interfaces that mediate the interactions between Bach2 and its co-repressors; it is not yet known whether the lateral groove of the Bach2 POZ domain is a critical site of such interactions and whether its co-repressor binding region thus resembles that of BCL6. X-ray crystal structures of the Bach2 POZ domain could then be used to develop inhibitors of co-repressor interactions using both virtual screening and rational design.
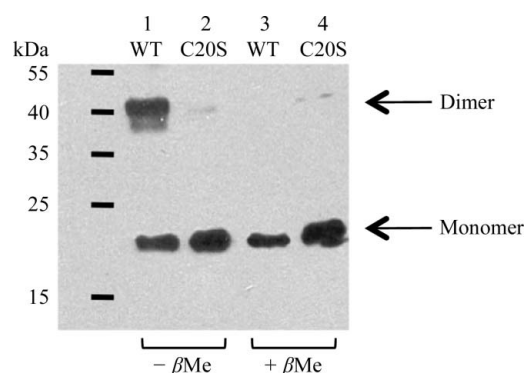
Figure 4
Oligomeric state of FLAG-tagged mouse Bach2 POZ domains expressed in COS-7 cells. Wild-type and C20S mutant FLAG-tagged Bach2 POZ domains were expressed in COS-7 cells and the proteins were analysed by SDS–PAGE under nonreducing (lanes 1 and 2) and reducing (lanes 3 and 4) conditions.

References

Ahmad, K. F., Engel, C. K. & Privé, G. G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12123–12128.

Ahmad, K. F., Melnick, A., Lax, S., Bouchard, D., Liu, J., Kiang, C.-L., Mayer, S., Takahashi, S., Licht, J. D. & Privé, G. G. (2003). *Mol. Cell*, **12**, 1551–1564.

Cerchietti, L. C. *et al.* (2010). *Cancer Cell*, **17**, 400–411.

Cerchietti, L. C., Yang, S. N., Shaknovich, R., Hatzi, K., Polo, J. M., Chadburn, A., Dowdy, S. F. & Melnick, A. (2009). *Blood*, **113**, 3397–3405.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Cohen, S. X., Ben Jelloul, M., Long, F., Vagin, A., Knipscheer, P., Lebbink, J., Sixma, T. K., Lamzin, V. S., Murshudov, G. N. & Perrakis, A. (2008). *Acta Cryst.* D**64**, 49–60.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Fox, G. C., Shafiq, M., Briggs, D. C., Knowles, P. P., Collister, M., Didmon, M. J., Makrantoni, V., Dickinson, R. J., Hanrahan, S., Totty, N., Stark, M. J., Keyse, S. M. & McDonald, N. Q. (2007). *Nature (London)*, **447**, 487–492.

Ghetu, A. F., Corcoran, C. M., Cerchietti, L., Bardwell, V. J., Melnick, A. & Privé, G. G. (2008). *Mol. Cell*, **29**, 384–391.

Hara, E., Takahashi, K., Tominaga, T., Kumabe, T., Kayama, T., Suzuki, H., Fujita, H., Yoshimoto, T., Shirato, K. & Shibahara, S. (1996). *Biochem. Biophys. Res. Commun.* **224**, 153–158.

Hoshino, H., Kobayashi, A., Yoshida, M., Kudo, N., Oyake, T., Motohashi, H., Hayashi, N., Yamamoto, M. & Igarashi, K. (2000). *J. Biol. Chem.* **275**, 15370–15376.

Hoshino, H., Nishino, T. G., Tashiro, S., Miyazaki, M., Ohmiya, Y., Igarashi, K., Horinouchi, S. & Yoshida, M. (2007). *J. Biochem.* **141**, 719–727.

Igarashi, K., Hoshino, H., Muto, A., Suwabe, N., Nishikawa, S., Nakauchi, H. & Yamamoto, M. (1998). *J. Biol. Chem.* **273**, 11783–11790.

Igarashi, K., Ochiai, K. & Muto, A. (2007). *J. Biochem.* **141**, 783–789.

Ito, N., Watanabe-Matsui, M., Igarashi, K. & Murayama, K. (2009). *Genes Cells*, **14**, 167–178.

Kamio, T., Toki, T., Kanezaki, R., Sasaki, S., Tandai, S., Terui, K., Ikebe, D., Igarashi, K. & Ito, E. (2003). *Blood*, **102**, 3317–3322.

Kobayashi, S., Taki, T., Chinen, Y., Tsutsumi, Y., Ohshiro, M., Kobayashi, T., Matsumoto, Y., Kuroda, J., Horiike, S., Nishida, K. & Taniwaki, M. (2011). *Genes Chromosomes Cancer*, **50**, 207–216.

Kobayashi, A., Yamagiwa, H., Hoshino, H., Muto, A., Sato, K., Morita, M., Hayashi, N., Yamamoto, M. & Igarashi, K. (2000). *Mol. Cell. Biol.* **20**, 1733–1746.

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. (2004). *Nucleic Acids Res.* **32**, W590–W594.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Motamed-Khorasani, A., Jurisica, I., Letarte, M., Shaw, P. A., Parkes, R. K., Zhang, X., Evangelou, A., Rosen, B., Murphy, K. J. & Brown, T. J. (2007). *Oncogene*, **26**, 198–214.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Muto, A., Hoshino, H., Madisen, L., Yanai, N., Obinata, M., Karasuyama, H., Hayashi, N., Nakauchi, H., Yamamoto, M., Groudine, M. & Igarashi, K. (1998). *EMBO J.* **17**, 5734–5743.

Muto, A., Ochiai, K., Kimura, Y., Itoh-Nakadai, A., Calame, K. L., Ikebe, D., Tashiro, S. & Igarashi, K. (2010). *EMBO J.* **29**, 4048–4061.

Muto, A., Tashiro, S., Nakajima, O., Hoshino, H., Takahashi, S., Sakoda, E., Ikebe, D., Yamamoto, M. & Igarashi, K. (2004). *Nature (London)*, **429**, 566–571.

Newberry, K. J., Fuangthong, M., Panmanee, W., Mongkolsuk, S. & Brennan, R. G. (2007). *Mol. Cell*, **28**, 652–664.

Ochiai, K., Muto, A., Tanaka, H., Takahashi, S. & Igarashi, K. (2008). *Int. Immunol.* **20**, 453–460.

Oyake, T., Itoh, K., Motohashi, H., Hayashi, N., Hoshino, H., Nishizawa, M., Yamamoto, M. & Igarashi, K. (1996). *Mol. Cell. Biol.* **16**, 6083–6095.

Painter, J. & Merritt, E. A. (2006). *J. Appl. Cryst.* **39**, 109–111.

Phan, R. T., Saito, M., Basso, K., Niu, H. & Dalla-Favera, R. (2005). *Nature Immunol.* **6**, 1054–1060.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A. & Noble, M. (2004). *Acta Cryst.* D**60**, 2288–2294.

Sanders, C. M., Sizov, D., Seavers, P. R., Ortiz-Lombardía, M. & Antson, A. A. (2007). *Nucleic Acids Res.* **35**, 3504–3515.

Schubot, F. D., Tropea, J. E. & Waugh, D. S. (2006). *Biochem. Biophys. Res. Commun.* **351**, 1–6.

Stead, M. A., Carr, S. B. & Wright, S. C. (2009). *Acta Cryst.* F**65**, 445–449.

Stead, M. A., Rosbrook, G. O., Hadden, J. M., Trinh, C. H., Carr, S. B. & Wright, S. C. (2008). *Acta Cryst.* F**64**, 1101–1104.

Stead, M. A., Trinh, C. H., Garnett, J. A., Carr, S. B., Baron, A. J., Edwards, T. A. & Wright, S. C. (2007). *J. Mol. Biol.* **373**, 820–826.

Stogios, P. J., Cuesta-Seijo, J. A., Chen, L., Pomroy, N. C. & Privé, G. G. (2010). *J. Mol. Biol.* **400**, 983–997.

Stogios, P. J., Downs, G. S., Jauhal, J. J., Nandra, S. K. & Privé, G. G. (2005). *Genome Biol.* **6**, R82.

Tashiro, S., Muto, A., Tanimoto, K., Tsuchiya, H., Suzuki, H., Hoshino, H., Yoshida, M., Walter, J. & Igarashi, K. (2004). *Mol. Cell. Biol.* **24**, 3473–3484.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.

Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. & Axel, R. (1979). *Cell*, **16**, 777–785.

Wood, M. J., Storz, G. & Tjandra, N. (2004). *Nature (London)*, **430**, 917–921.

Yoshida, C., Yoshida, F., Sears, D. E., Hart, S. M., Ikebe, D., Muto, A., Basu, S., Igarashi, K. & Melo, J. V. (2007). *Blood*, **109**, 1211–1219.

Zhang, K.-Y., Cowtan, K. & Main, P. (1997). *Methods Enzymol.* **277**, 53–64.

Zhou, P., Kalakonda, N. & Comenzo, R. L. (2005). *Br. J. Haematol.* **128**, 636–644.