JMB

ELSEVIER

Insight into Molecular Interactions Between Two PB1 Domains

Audrey van Drogen-Petit[1,2], Catherine Zwahlen[3], Matthias Peter[2]* and Alexandre M. J. J. Bonvin[4]*

[1]*Institute of Chemistry Swiss Federal Institut of Technology Lausanne (EPFL) BCH CH-1015 Lausanne Switzerland*

[2]*Institute of Biochemistry Swiss Federal Institut of Technology Zurich (ETHZ) ETH Hoenggerberg HPM G 6.2, CH-8093 Zürich Switzerland*

[3]*Geneprot, 2 Pré-de-la-Fontaine, CH-1217 Genève, Switzerland*

[4]*Bijvoet Center for Biomolecular Research Utrecht University Padualaan 8, 3584CH Utrecht The Netherlands*

Specific protein–protein interactions play crucial roles in the regulation of any biological process. Recently, a new protein–protein interaction domain termed PB1 (Phox and Bem1) was identified, which is conserved throughout evolution and present in diverse proteins functioning in signal transduction, cell polarity and survival. Here, we investigated the structure and molecular interactions of the PB1 heterodimer complex composed of the PB1 domains of the yeast proteins Bem1 and Cdc24. A structural model of the Cdc24 PB1 was built by homology modeling and molecular dynamics simulations, and experimentally validated by ^{15}N nuclear Overhauser effect spectroscopy (NOESY)–heteronuclear single quantum coherence (HSQC) analysis. Residues at the interface of the complex for both proteins were identified by NMR titration experiments. A model of the heterodimer was obtained by docking of the two PB1 domains with HADDOCK, which applies ambiguous interaction restraints on residues at the interface to drive the docking procedure. The refined model was validated by site-directed mutagenesis on both Bem1 and Cdc24. Finally, the docking was repeated from the newly published NMR structure of Cdc24, allowing us to assess the performance of homology-based docking. Our results provide insight into the molecular structure of the Bem1–Cdc24 PB1-mediated heterodimer, which allowed identification of critical residues at the binding interface.

© 2004 Elsevier Ltd. All rights reserved.

Corresponding authors

Keywords: PB1 domains; docking; HADDOCK; Cdc24; Bem1

Introduction

Specific protein interaction domains play essential roles in the signaling output of a wide range of cell-surface receptors and are, in turn, controlling numerous cellular events, such as protein and vesicle trafficking, cell-cycle progression, regulation of gene expression, DNA replication and repair, cytoskeletal rearrangements and targeted protein degradation.[1] Many specific protein–protein interaction domains have been discovered, including SH2, PX, PH and SH3 motifs. In addition, several domains mediate specific binding to small molecules such as phospholipids and sugars, or modified amino acids like phosphotyrosine and phosphoserine. Current knowledge regarding these interactions has been obtained mainly *via* mutational analysis and, in most cases, the specific molecular determinants that contribute to the specificity of these interactions remain unclear.

Recently, a new protein–protein interaction domain has been characterized, termed the PB1 domain for Phox and Bem1.[2,3] Besides Phox and fungal Bem1 proteins, the PB1 motif is found in several signaling proteins including human MEK3, several isoforms of protein kinase C and numerous hypothetical proteins of *Arabidopsis thaliana*.[3] Pairs of PB1 domains form specific heterodimers and have been shown to mediate

functional protein–protein interactions in a variety of biological processes. For example, the adaptor protein Bem1 and the guanine nucleotide exchange factor (GEF) Cdc24 interact through their PB1 domains and initiate the establishment of cell polarity in budding yeast by regulating the assembly of the actin cytoskeleton.[4] Likewise, p40phox/ p67phox interact through PB1 domains and are part of the oxidase complex that is required to kill pathogens in the phagosome.[5] Finally, the PKCζ/ ZIP PB1-mediated heterodimer is required to recruit the kinase into a signaling complex containing potassium channels.[6,7] The structure of the PB1 domain of Bem1 has been determined recently by nuclear magnetic resonance spectroscopy (NMR).[8] It showed a β-grasp fold similar to Ras^{GTP}-binding domains (SCOP classification).[9] Recently, the first structure of a PB1 complex, the PB1 mammalian pair p67phox/p40phox, has been characterized by X-ray crystallography,[10] and a model of the ZIP/ λPKC PB1-mediated heterodimer has been proposed from comparative modeling and docking.[11] Most PB1 domains contain a highly conserved motif termed OPR (octicosapeptide repeat) that is required for the formation of heterodimers such as Bem1/Cdc24, p40phox/p67phox, ZIP/λPKC.[2,5,11] Key residues for the interaction of PB1 domains have been identified by mutagenesis studies. For example, mutations of K480 or K482 in the Bem1 PB1 domain abolish its binding to the Cdc24 PB1 domain both *in vivo* and *in vitro*.[3,12] In the case of the mammalian pair p40phox/p67phox, an alanine scanning mutagenesis study determined that the p40phox OPR motif is essential for binding to the p67phox PB1 domain.[5] Further studies stress the importance of acidic clusters embedded in the OPR motif, as well as basic residues that are part of the OPR-binding motif.[10,11] Indeed, for the p67phox/p40phox and ZIP/λPKC PB1 containing proteins, such basic and acidic clusters seem to establish coulombic interactions at the interface of the complex and contribute to the specificity of the PB1–PB1 heterodimers observed *in vivo* and *in vitro*.[10] At the date of submission, no three-dimensional structure of the PB1 domain of Cdc24 (780–854) has been reported, despite a previous NMR study reporting its secondary structure.[8]

Here, we investigated the structure and molecular interactions of the PB1 heterodimer complex composed of the PB1 domains of Bem1 and Cdc24. A structural model of the Cdc24 PB1 was built by homology modeling and refined using molecular dynamics (MD) simulations. The MD-refined model, which underwent a significant rearrangement of a strand, was validated by the observation of a few crucial amide–amide nuclear Overhauser effect (NOEs) contacts, in a ^{15}N NOE spectroscopy (NOESY)–heteronuclear single quantum coherence (HSQC) NMR experiment. Residues at the interface of the complex for both proteins were identified by NMR titration experiments that allow to monitor complex formation at a residue level. A model of the heterodimer was obtained

by docking of the two PB1 domains with HADDOCK, which applies ambiguous interaction restraints on residues at the interface to drive the docking procedure.[13] The model obtained was validated by site-directed mutagenesis on both Bem1 and Cdc24. Finally, since the NMR structure of Cdc24 PB1[37] became available after initial submission of this work, the docking was repeated, based on the experimental structure of the longer Cdc24 construct (757–854), which allows us to assess the performance of the homology-based docking.

Results

Identification of binding sites for the Bem1 PB1/Cdc24 PB1 complex

Molecular interactions between Bem1 PB1 and Cdc24 PB1 were characterized by NMR titration experiments using ^{1}H–^{15}N HSQC spectra. Two sets of titrations were performed: (i) HSQC spectra of labeled ^{15}N-Bem1 PB1 were recorded in the absence or in the presence of increasing amounts of unlabeled Cdc24 PB1 (the molar ratio of Cdc24 PB1/Bem1 PB1 varied between 0.16 and 1); and (ii) the ^{1}H and ^{15}N chemical shifts of ^{15}N-Cdc24 PB1 were monitored upon gradual addition of unlabeled Bem1 PB1. During both titrations (^{15}N-Cdc24 PB1/ Bem1 PB1; ^{15}N-Bem1 PB1/ Cdc24 PB1), we observed only one set of resonances and could follow the shifting peaks, meaning that the complex is in fast exchange on the NMR time-scale. The combined (^{1}H/^{15}N) chemical shift perturbations for both domains are shown in Figure 1. For the ^{15}N-Bem1 PB1/Cdc24 PB1 complex (Figure 1A), 16 amino acid residues of Bem1 PB1 show significant combined chemical shift perturbations upon complex formation ((Δppm > 0.009), see Materials and Methods); namely, K477, I481, K482, F483, Y484, Y485, D488, F490, K506, I507, R510, I511, N515, S547, D550 and I551. Those residues were mapped onto the structure of the domain (PDB code 1ipg), and cluster on a well-defined surface of the protein as shown in Figure 1A. From the NMR titration of the complex Bem1 PB1/^{15}N-Cdc24 PB1, 18 residues of Cdc24 PB1 were identified that show significant (at least one standard deviation above average) chemical shift perturbations (Δppm > 0.001): S778, F781, K787, W789, I802, N804, K817, Y818, L828, G829, E832, W834, N835, E843, N844, N845, E846, and F848 (Figure 1B). Interestingly, the two groups of residues contain mainly charged and aromatic amino acid residues, suggesting that the binding between the two domains is possibly mediated by electrostatic interactions and aromatic stacking.

Determination and validation of Cdc24 PB1 structure model

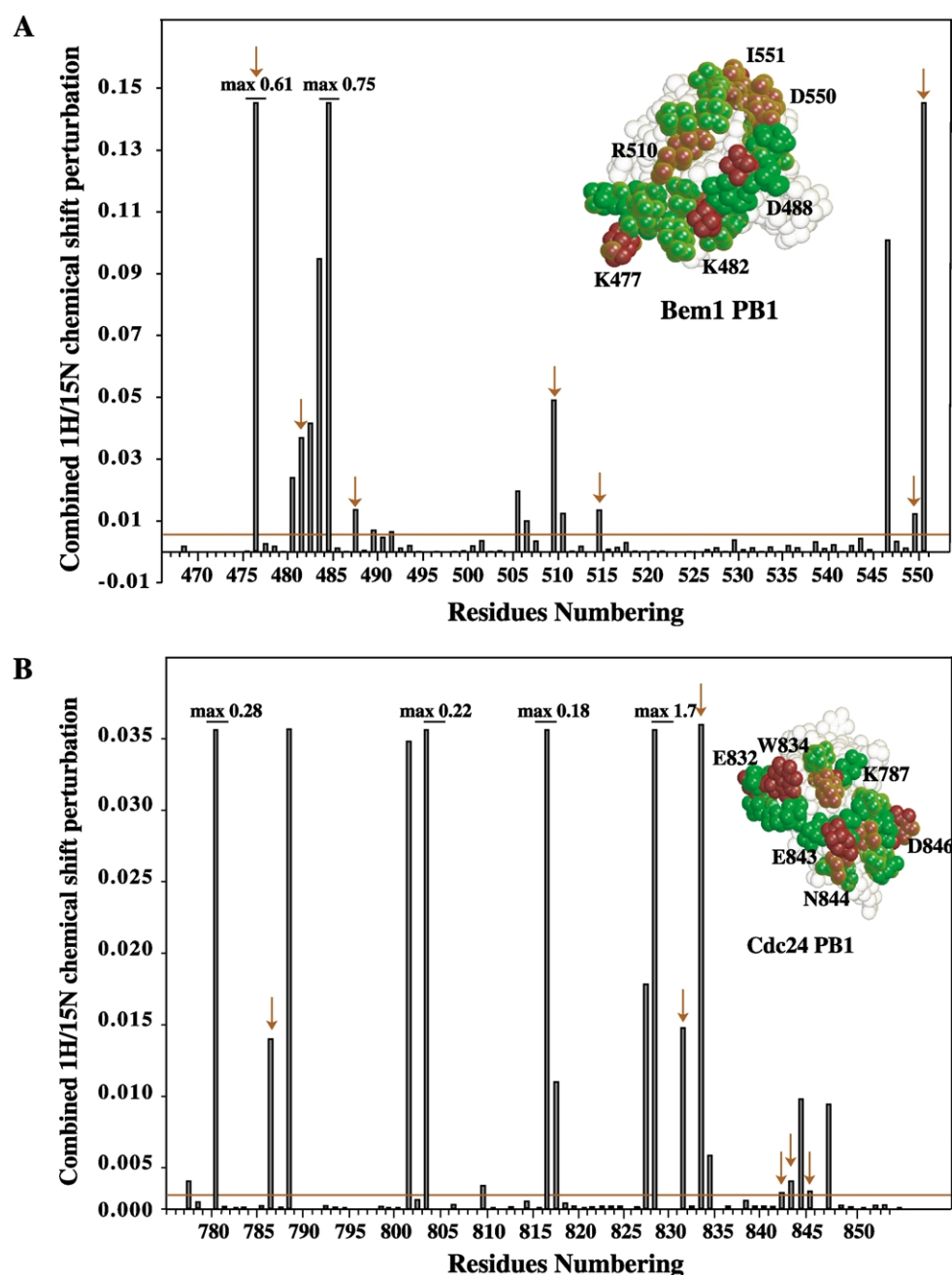Since no structure of Cdc24 PB1 was available,

Figure 1. Combined amide (^{1}H) and nitrogen (^{15}N) chemical shift perturbations ($[(\Delta\delta^1H_N)^2 + (\Delta\delta^{15}N_{ppm}/6.51)^2]^{0.5}$) for Bem1 PB1 and Cdc24 PB1 as a function of the amino acid sequence (refer to Figure 5 for numbering). A, Chemical shift changes of ^{15}N-Bem1 PB1 upon binding to unlabeled Cdc24 PB1. B, Chemical shift changes of ^{15}N-Cdc24 PB1 upon binding to unlabeled Bem1 PB1. The horizontal line represents the calculated average chemical shift perturbation and the arrows point out active residues selected as ambiguous interaction restraints for the docking. The amino acid residues defined as being part of the binding site are displayed at the top right of each graph on the Bem1 PB1 3D structure and on the Cdc24 PB1 3D model. These are colored following the definition used for docking: in red the active residues (with significant chemical shift perturbation and solvent-accessible) and in green the passive residues (surface neighbors of active residues).

we built a model for Cdc24 (780–854) that was subsequently used for the docking to Bem1 PB1. To establish the model of Cdc24 PB1, we first searched for a suitable template in PDB[14] using threading techniques with the program PROSPECT.[15] The best candidate was the Ras-associating domain of the rat protein Ral GEF (PDB entry 1lfd, chain A) with a sequence identity of 19% (Figure 5B). This domain was used to generate a homology 3D model of Cdc24 PB1 with the program Modeller 4.[16] The resulting model (Figure 2A) comprises two α helices (from residues 791 to 803 and residues 837 to 841), and three anti-parallel β sheets (from residues 813 to
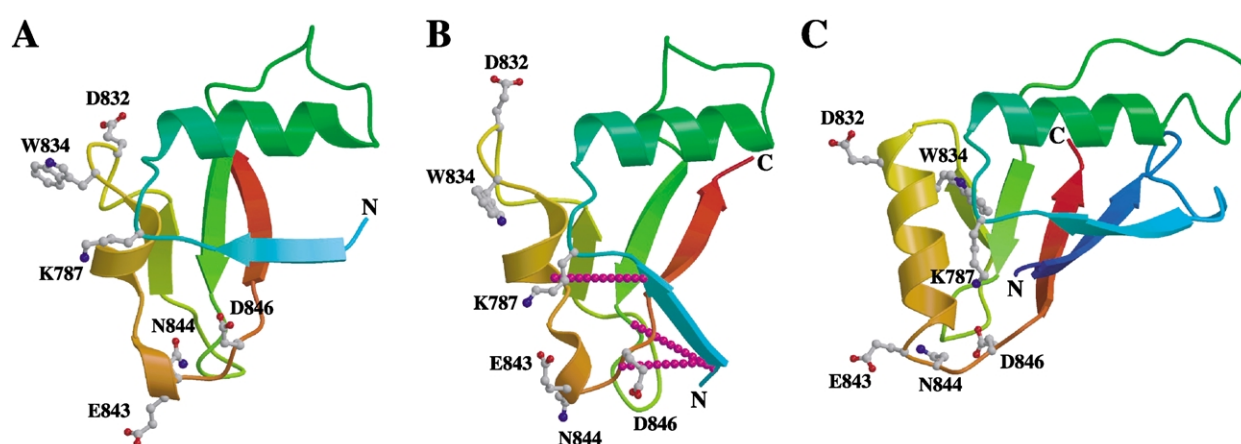
Figure 2. Structural model of the PB1 domain of Cdc24 and comparison with the NMR solution structure. A, Cdc24 PB1 (780–854) 3D homology model based on sequence alignment with the Ras-associating domain of GNDS_RAT and generated with Modeller 4.[16] B, Snapshot taken at 15 ns from the MD simulation (see Materials and Methods) and (C) recently published NMR solution structure (757–854).[23] C, The NMR structure contains 19 extra residues at the N terminus, forming an additional β-strand. The structures are colored from blue, N terminus, to red, C terminus. The pink lines in B indicate a few crucial HN-HN NOEs (F781-I810, S782-E846, T785-M840) extracted from the ^{15}N-NOESY-HSQC spectrum that validate the rearrangement of the N-terminal strand in the model generated by MD simulation. Side-chains of residues involved in binding to Bem1 identified from NMR titrations (see the text) are shown for reference. The Figures were generated using MOLSCRIPT[37] and Raster3D.[38]

819, residues 825 to 827 and residues 848 to 853). A comparison of our model with the Ras template revealed that our sequence is missing an N-terminal fragment that is sandwiched between the N and C-terminal strands in the Ras-associating domain, thereby forming a triple-stranded β-sheet. Our sequence corresponds, however, to the shortest region identified as necessary and sufficient for tight binding to Bem1 PB1,[2,3,8,12] indicating that the missing part is not interacting with Bem1. To assess the effect of this missing fragment on the 3D structure, the validity and stability of our homology model was investigated by extensive MD simulations in explicit solvent. The solvated protein was simulated for 10 ns at constant temperature (300 K) and pressure (1 atm = 101,325 Pa) under periodic boundary conditions. The secondary structure elements present in the initial model are well maintained throughout the 10 ns simulation (data not shown). A structural transition takes place, however, after approximately 5 ns, as can be observed clearly from the plot of the pairwise positional root-mean-square deviation (RMSD) matrix in Figure 3. This conformational change involves a rearrangement of the N-terminal residues (780–791) that move to pack against the C-terminal strand. The resulting conformation, shown in Figure 2B is stable throughout the remainder of the MD simulation (large off-diagonal blue block after 5 ns in Figure 3). An analysis of the structure at this point indicated a number of close amide–amide contacts between N and C-terminal residues that should be easy to detect by NMR. Indeed, these interactions were confirmed by analysis of the 3D ^{15}N-NOESY-HSQC spectrum of Cdc24 PB1. 24 amide-amide NOEs could be assigned (Table 1s of the Supplementary Material). Three amide–

amide NOEs were identified between the N-terminal and C-terminal residues, T782-E846, L783-L849 and V785-M840, which confirmed the model obtained after the conformational transition observed during the MD simulation. The remaining amide–amide NOEs confirm the secondary structural elements in our model, which are in accordance with the prediction by Terasawa et al. (Figure 5F).[8]

The assigned NOEs (Table 1s of the Supplementary Material) were introduced as distance restraints and the MD in explicit solvent was extended by 5 ns to a total of 15 ns. Inclusion of the experimental restraints did not result in any significant changes in the model (Figure 3). These results validate our MD-refined homology model, which we used in the following to generate a model of the complex.

The NMR structure of a longer Cdc24 construct (757–854), which became available after submission of this work (PDB entry 1Q1O),[17] enabled us to assess the quality of our homology model. Its structure is presented in Figure 2C for comparison. Considering the low level of sequence identity (19%) of the template, the Ras-associating domain, which we used for homology modeling, the quality of our model is quite remarkable (compare Figure 2B and C). If one neglects the N-terminal strand that is rearranged in our model due to the missing N-terminal fragment (757–779) in agreement with the NOE data, the overall fold is well predicted. The RMSD between our model and the NMR structure for segments 791–802 and 816–852 (i.e. neglecting the rearranging N-terminal strand and the disordered region 803–816 in the NMR structure) is 4.5 Å for backbone atoms. If one only considers one the two helices, the RMSD
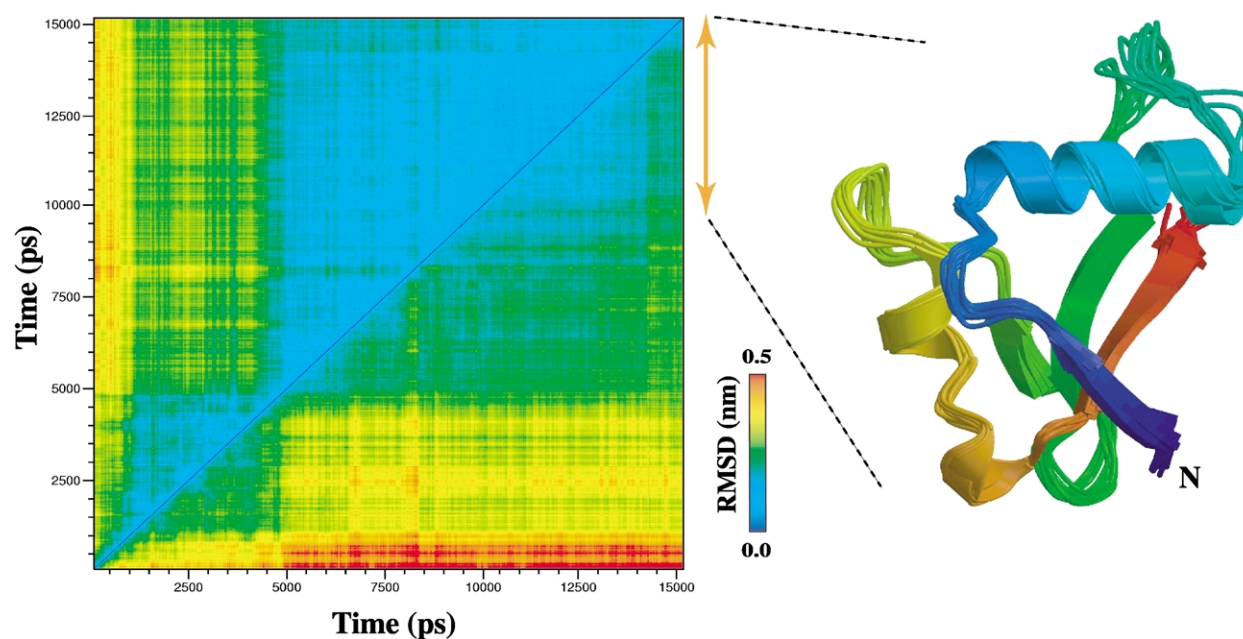
Figure 3. Pairwise RMSD matrix of Cdc24 PB1 as a function of the MD simulation time. Upper left triangle: backbone RMSD, lower right triangle: all heavy-atom RMSD. The RMSD values are color-coded from blue (0 nm) to red (0.5 nm). The ensemble of structures on the right corresponds to snapshots of the MD simulation taken every 0.5 ns between 10 ns and 15 ns. The structure reached a stable conformation after approximately 5 ns simulation, as indicated by the large blue off-diagonal block.

value decreases to 3.0 Å. The most striking difference concerns the second helix, which is poorly formed in our model. Most residues interacting with Bem1 as identified from the NMR titrations are located in this second helix. The RMSD values between our model and the NMR structure for the active residues identified as part of the interface (see below) are 2.7 Å and 4.6 Å for backbone and all heavy-atoms, respectively. As we will see in the following, this results in a difference in the relative orientation of the two proteins in the complex between the homology and NMR-based docking solutions. The interacting surfaces are unaffected to a major extent.

Homology-based docking of Bem1 PB1 and Cdc24 PB1 using HADDOCK

A 3D model of the PB1 heterodimer was built by docking using HADDOCK.[13] The docking was started from two ensembles of structures, (i) the solution structure of Bem1 PB1 (PDB entry 1ipg) containing 20 models and (ii) ten conformations of Cdc24 PB1 taken every 0.5 ns from the last 5 ns of the MD simulation (10.5–15 ns). In HADDOCK, the docking is driven by ambiguous interaction restraints (AIRs) that are defined between residues showing significant chemical shift perturbations in the NMR titration experiments and that are solvent-accessible (see Materials and Methods). For Bem1 PB1, seven residues were defined as active (K477, K482, D488, R510, N515, D550 and I551), in addition, 13 solvent-accessible surface neighbors were introduced as passive residues

(T478, K480, K486, D487, I489, A491, M493, K495, P509, D512, D514, K545 and H549). For Cdc24 PB1, the binding interface was defined by six active residues (D787, E832, W834, E843, N844 and E846) and by 11 passive residues (T782, L784, E786, V788, Q819, D831, D833, K838, L841, A842 and K847). This resulted in a total of 13 AIRs defined from each active residue of one partner to all active and passive residues of the binding partner (Table 1).[13]

For the docking, 1600 rigid-body docking structures were first generated and the best 200 in terms of intermolecular interaction energy ($E_{vdw} + E_{elec} + E_{AIR}$) were further subjected to a semi-flexible simulated annealing and a final refinement in water. The flexible interface comprised segments K477-D497, I507-K517 and S534-I551 for Bem1 PB1 and I780-N790, K817-E821 and E821-L849 for Cdc24. Figure 4A shows the intermolecular energy as a function of the backbone RMSD from the lowest-energy structure for the 200 structures obtained after water refinement. Four clusters could be identified based on pairwise backbone RMSD values using a 1.5 Å cut-off. The lowest-energy cluster ($E_{intermolecular} = -791(\pm 130)$ kcal/mol) containing 23 members satisfies best the AIR restraints and shows the largest buried surface area at the interface. The ten lowest-energy structures of this cluster (Figure 4C) were selected for further analysis.

The structure of the complex is rather well defined, considering the limited amount of experimental information, with an overall backbone RMSD from the mean structure of 1.1(± 0.6) Å (see

Table 1. List of active and passive residues used in the definition of the ambiguous interaction restraints (AIRs) for docking of Bem1 and Cdc24-PB1, and flexible segments

Bem1	Active residues[a]	K477, K482, D488, R510, N515, D550, I551
	Passive residues[b]	T478, K480, K486, D487, I489, A491, M493, K495, P509, D512, D514, K545, H549
	Flexible segments	K477-D497, I507-K517, S534-I551
Cdc24 (780–854) (homology model)	Active residues[a]	D787, E832, W834, E843, N844, E846
	Passive residues[b]	T782, L784, E786, V788, Q819, D831, D833, K838, L841, A842, K847
	Flexible segments	I780-N790, K817-E821, E821-L849
Cdc24 (757–854) (NMR)	Active residues[a]	D787, E832, N835, E843, N844, N845
	Passive residues[b]	G760, E786, V788, N790, Q819, E821, D822, G823, S830, D831, D833, V836, E839, A842, K847, F848, N850
	Flexible segments	G757 -I762, L784-D792, K817-F825, L828-L849

For Cdc24, the residues are listed for both the homology model and the NMR solution structure. Differences between the two originate from differences in solvent-accessibility.
[a] The active residues correspond to the residues having a significant NMR chemical shift perturbation during the NMR titration experiments and that are highly solvent-accessible.
[b] The passive residues correspond to all surface neighbors of the active residues that are solvent-accessible.

structural statistics in Table 2) and of good quality with more than 96% of the residues in the most favored and allowed regions of the Ramachandran plot. The complex interface buries 1437 Å² of surface area and, as can be seen from the inter-molecular energies in Table 2, is dominated by electrostatic interactions. A large number of inter-molecular hydrogen bonds and salt-bridges could be identified, together with a few hydrophobic contacts, the major one between Y485 in Bem1 and L841 in Cdc24 (Table 3).

A comparison of the identified intermolecular interactions (Table 3) with multiple sequence alignments revealed that for Bem1 PB1, 11 of the 14 residues involved in intermolecular hydrogen bonds or hydrophobic contacts are highly conserved throughout the species (Figure 5). For Cdc24 PB1, ten of the 14 residues involved in intermolecular interactions are located in a known conserved motif called the OPR motif.[18] Particularly note-worthy are the interactions observed between K480[Bem1 PB1] and E832[Cdc24 PB1], K480[Bem1 PB1] and D833[Cdc24 PB1] and between K482[Bem1 PB1] and D833[Cdc24 PB1]. The two lysine residues (480 and 482) in Bem1 PB1 are known to be essential for the complex formation: mutation of these two key residues abolishes the complex formation.[3,12] In Cdc24 PB1, D833 remained conserved throughout evolution (Figure 5B). These results clearly stress the importance of the interactions involving these residues (K480[Bem1 PB1], K482[Bem1 PB1], E832[Cdc24 PB1] and D833[Cdc24 PB1]) at the interface of the complex. In order to validate our model and the proposed interactions, we mutated K482[Bem1 PB1], E832[Cdc24 PB1] and D833[Cdc24 PB1] swapping their identities between Bem1 and Cdc24. Our proposed model predicts that the single mutations resulting in a change of charge should abolish binding, while the double, altered-specificity mutants should be able to restore binding. The mutant proteins were purified from *Escherichia coli*, and the Cdc24 PB1 mutant was tested for its ability to bind to the gluta-thione-*S*-transferase (GST)-Bem1 PB1 immobilized on GST-beads. As expected, both single mutants were defective for their interaction with wild-type Cdc24 PB1 or Bem1 PB1, respectively (Figure 6). Strikingly, however, the Bem1 PB1-K482E was able to bind to Cdc24 PB1-E832K-D833K, demonstrating that the amino acid residues E832-D833[Cdc24 PB1] and K482[Bem1 PB1] interact with each other. Taken together, the mutagenesis and swapping experiments strongly support the proposed model of the molecular interactions between the two PB1 domains in the heterodimer.

NMR-based docking of Bem1 PB1 and Cdc24 PB1 using HADDOCK

The publication of the NMR structure of Cdc24 (757–854)[17] (PDB entry 1Q1O) allowed us to repeat the docking to control if the malformed second helix in our homology model leads to differences at the Bem1–Cdc24 interface. The difference in solvent-accessibilities between the NMR structure and the homology model results in two differences in the definition of active residues for Cdc24: W834 is replaced by N835 and E846 by N845. The definition of passive residues and of the flexible segments for docking is affected (see Table 1). The number of AIRs, however, remains the same (13). The docking was performed in the same manner as that used with the homology model. Again, four clusters are obtained, with the lowest cluster being the most populated (125 members) (Figure 4B). The proper definition of the second helix clearly leads to a clearer distinction of docking solutions. The ten lowest-energy structures of the best cluster are shown in Figure 4D. The Bem1–Cdc24 interface is quite similar to the solution obtained with the homology model. The most striking difference between the two is an approximate 43° rotation of Cdc24 around the axis defined by the second helix. This originates from the different orientations of the active residues due to the malformed second helix in the homology model. Despite this, the interfaces are very similar,
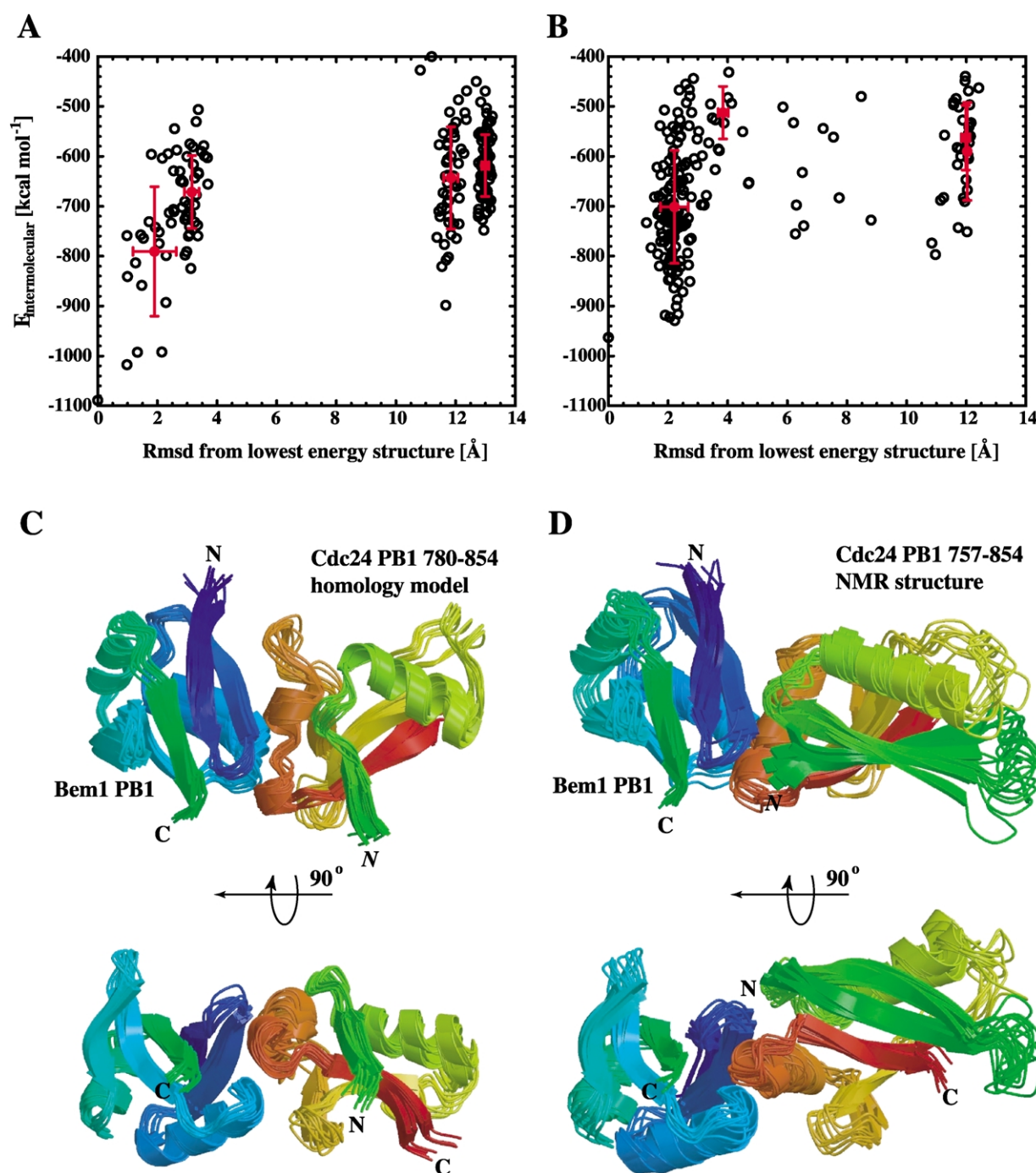
Figure 4. HADDOCK docking solutions for the Bem1/Cdc24 PB1 complex based on the homology model (A and C) and the NMR structure (B and D) of Cdc24 PB1. A and B, Intermolecular energies as a function of the backbone RMSD on the flexible interface (see Table 1). C, Ensemble of the ten lowest-energy structures of the lowest-energy cluster of the homology-based docking. D, Ensemble of the ten lowest-energy structures of the lowest-energy cluster of the NMR-based docking. A definition of the residues identified from NMR titration experiments and used to define ambiguous interaction restraints can be found in Table 1. The structure representations were generated with the programs MOLSCRIPT[37] and Raster3D.[38]

with 92% and 77% of the interface residues of Bem1 and Cdc24, respectively, shared between the homology-based and NMR-based Bem1-Cdc24 models. The exact identity of intermolecular contacts (Table 3) is, however, more affected by the difference in the second helix, with 42% of the contacts (eight out of 19) conserved between the two

solutions (see Table 3). These conserved contacts span the entire surface of the interface, resulting in a similar binding mode, which clearly validates our homology-based docking. A difference between the homology-based and NMR-based Bem1-Cdc24 models should be noted in the contacts that K480$^{Bem1\ PB1}$ and K482$^{Bem1\ PB1}$ are making

Table 2. Structural statistics of the ten best Bem1–Cdc24 model structures from homology and NMR-based docking

	Bem1–Cdc24	
	Homology-based	NMR-based
Backbone RMSD (Å) with respect to mean[a]		
Flexible interface backbone	0.90 ± 0.40	1.20 ± 0.50
All backbone	1.05 ± 0.55	1.50 ± 0.75
Number of ambiguous interaction restraints (AIRs)		
From Bem1	7	7
From Cdc24	6	6
Total AIRs	13	13
CNS intermolecular energies after water refinement[b]		
E_{vdw} (kcal mol^{-1})	-37 ± 9	-34 ± 6
E_{elec} (kcal mol^{-1})	-913 ± 102	-875 ± 29
Buried surface area (Å^2)[c]	1437 ± 69	1362 ± 84
RMSD from idealized covalent geometry		
Bonds (Å)	0.003 ± 0.00	0.003 ± 0.00
Angles (deg.)	0.42 ± 0.01	0.39 ± 0.01
Impropers (deg.)	0.36 ± 0.01	0.34 ± 0.01
Ramachandran analysis		
Residues in the favored region (%)	71.7	82.2
Residues in additionally allowed regions (%)	24.9	15.3
Residues in generously allowed regions (%)	1.1	0.9
Residues in disallowed regions (%)	2.3	1.7

Based on ten structures obtained after flexible docking with HADDOCK followed by refinement in explicit water using ambiguous interaction restraints derived from chemical shift perturbation data (see Materials and Methods). The docking was first performed with the homology model of Cdc24 PB1 and repeated after initial submission once the NMR structure of Cdc24 became available.[17]

[a] See definition of flexible interface in Table 1. For comparison, the backbone RMSDs of the three ensembles of structures used for docking were 0.32 Å for Bem1 (477–551) (20 structures) and 0.74 Å and 1.15 Å for Cdc24 homology model (780–854) (ten structures) and Cdc24 NMR (757–854) (20 structures), respectively.

[b] The non-bonded energies were calculated with the OPLS parameters using an 8.5 Å cut-off.

[c] The buried surface area was calculated with CNS using a 1.4 Å radius water probe and 0.05 Å accuracy.

with Cdc24: instead of contacting mainly E832$^{Cdc24\ PB1}$ and D833$^{Cdc24\ PB1}$, they contact D831$^{Cdc24\ PB1}$ and E832$^{Cdc24\ PB1}$ in the NMR-based model. D833$^{Cdc24\ PB1}$ is no longer involved in intermolecular contacts. For validation (see above), we chose to mutate both E832$^{Cdc24\ PB1}$ and D833$^{Cdc24\ PB1}$. The former is clearly important, since it contacts both K480$^{Bem1\ PB1}$ and K482$^{Bem1\ PB1}$, which explains why our mutagenesis and swapping experiments were successful. A detailed view of the interface in the Bem1-Cdc24 model based on the NMR structures of both proteins, which we now consider as our final model, is shown in Figure 7.

Discussion

In this study, we investigated the structure of the PB1 domain of Cdc24 by modeling and MD simulations. The model was later validated by comparison with the newly published NMR structure of Cdc24. The binding interface of the PB1 heterodimer of Cdc24 and Bem1 was characterized by NMR titrations and docking. Several key residues that mediate specific binding of the two domains were identified and these interactions were confirmed by mutational analysis. Our results allow us to propose a model for the specific interactions of PB1 domains.

The role of the Cdc24–Bem1 interaction *in vivo*

Cdc24 and Bem1 are involved in the establishment of cell polarity in the budding yeast *Saccharomyces cerevisiae*.[19] Binding of Cdc24 to Bem1 at the cell cortex stabilizes active Cdc24, and leads to the sustained activation of the small GTPase Cdc42 at sites of polarized growth.[20] Bem1 mutants unable to interact with Cdc24 polarize inefficiently, because Cdc24 cannot be maintained at growth sites.[12] The interaction between Cdc24 and Bem1 is mediated through their PB1 domains. Phosphorylation of Cdc24 by the PAK-like kinase Cla4 is thought to trigger dissociation of Bem1 and Cdc24.[20] Although the sites of phosphorylation have not been mapped, it is tempting to speculate that they may be contained within the PB1 domains of the two proteins. Interestingly, several serine and threonine residues are located in the interface between Cdc24 PB1 (S830 and T782) and Bem1 PB1 (T478 and T505), which, if phosphorylated, could likely disrupt binding.

Model of Cdc24 PB1

An approach combining homology modeling from an alignment obtained by threading and MD simulation was chosen to generate the Cdc24 PB1 model. The model was validated by NMR with the observation of a few crucial NOEs in a

Table 3. Intermolecular contacts statistics calculated over the ensemble of the ten best structures of the two HADDOCK runs based on the homology model or the NMR structure of Cdc24 (see the text)

Donor			Acceptor		Hydrogen-bonds		Non-bonded contacts
					S–S	M–S	
A. Common contacts for both runs[a]							
Lys480	A	→	Asp831	B	4/4	0	0
Lys480	A	→	Glu832	B	6/9	1/0	0
Lys486	A	→	Glu843	B	5/7	4/0	9/0
Phe490	A	→	Glu839	B	0	0	5/9
Arg510	A	→	Glu839	B	10/7	0	5/7
Lys787	B	→	Asp487	A	10/9	1/2	0/5
Lys787	B	→	Asp488	A	7/7	0	1/6
Lys838	B	→	Asp488	A	1/9	0	3/5
B. Bem1–Cdc24 (homology model) specific contacts							
Lys480	A	→	Asp833	B	6	1	4
Lys482	A	→	Asp833	B	6	0	0
Tyr485	A	→	Leu841	B	0	0	10
Arg510	A	→	Phe825	B	0	0	7
Arg510	A	→	Asp831	B	6	0	0
Lys817	B	→	Asp512	A	9	0	1
Gln819	B	→	Asp512	A	6	1	2
Lys838	B	→	Ile489	A	0	7	4
Met840	B	→	Asp488	A	0	7	0
Leu841	B	→	Asp488	A	0	7	7
Ala842	B	→	Asp487	A	0	6	1
Glu843	B	→	Asp487	A	0	5	6
Asn844	B	→	Tyr485	A	9	0	0
Asn845	B	→	Asp512	A	7	0	5
C. Bem1–Cdc24 (NMR structure) specific contacts							
Lys482	A	→	Asp831	B	9	0	0
Lys482	A	→	Glu832	B	8	0	0
Lys482	A	→	Asn835	B	7	0	0
Tyr485	A	→	Glu843	B	7	0	7
Ala491	A	→	Glu839	B	0	5	0
Arg510	A	→	Glu832	B	8	2	0
Arg510	A	→	Val836	B	0	0	6
Glu843	B	→	Thr513	A	1	0	7
Asn844	B	→	Asp512	A	6	1	0
Lys847	B	→	Asp487	A	9	0	2
Lys847	B	→	Asp488	A	7	0	0

Intermolecular contacts were analyzed with DIMPLOT[36] (see Materials and Methods) and are reported if present in at least five of the ten best structures. The occurrence of side-chain–side-chain (S–S) and side-chain-main–chain (S–M) hydrogen bonds is reported. Chain A corresponds to Bem1 and chain B to Cdc24. The number of occurrences of a given interaction over the ensemble of ten best structures is reported.
[a] The occurrence in each run is reported as run Cdc24 homology model/run Cdc24 NMR.

3D-NOESY-HSQC experiment and is consistent with the secondary structure mapping from previous NMR studies.[8] Since Bem1 PB1 and Cdc24 PB1 have been found to be homologous domains,[9] we compared the fold of our Cdc24 PB1 model with that determined for Bem1 PB1. They appear to be similar although, as pointed out previously, Cdc24 PB1 has a missing strand at its N terminus (760–780). This "shorter" construct has been studied extensively and defined as the minimal construct necessary and sufficient for tight binding to the Bem1 PB1 domain.[2,8,12] Interestingly, as shown by MD simulation, the overall fold is stable even without the missing N terminus, once the repacking of the N-terminal strand has occurred. The packing deficiency originally introduced by this missing segment in our homology model is compensated by a repacking of the N-terminal strand against the C terminus, resulting in a more compact conformation (Figure 2B). In accordance with the previous experimental studies,[2,8,12] we can conclude that the 20 N-terminal amino acid residues are not essential for the binding to Bem1 PB1, and do not seem to be required to stabilize the 3D conformation of the Cdc24 PB1 domain. When comparing the homology model with the NMR structure of the longer Cdc24 construct, the overall improvement of the model after MD simulation is reflected only slightly in the backbone RMSD that decreases from 4.75 Å to 4.6 Å (neglecting the rearranging N-terminal strand and the disordered region 803–816 in the NMR structure). The second helix is clearly improving, however, as can be seen in Figure 2A and B. This is one of a few examples where MD simulations successfully and significantly improved the quality of a homology model. While, in the presence of a high level of sequence homology, it is generally believed
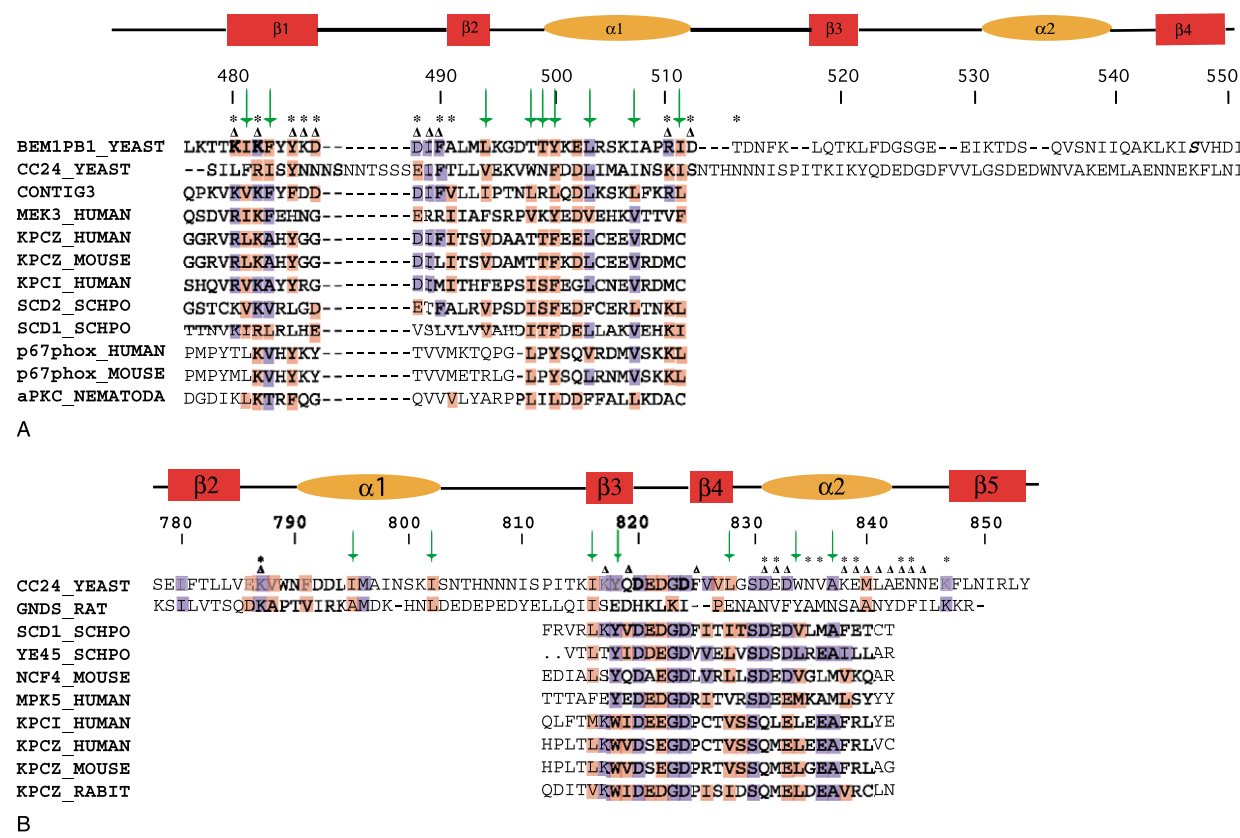
Figure 5. Sequence alignments of Bem1 PB1 (A) and Cdc24 PB1 domains (B). The numbering in Bem1 and Cdc24 starts from the amino-terminal methionine residue. Similar residues are highlighted in red and identical residues are highlighted in blue. The mutations previously characterized by Butty *et al.*[12] (Bem1: m1 K480E and m2 S547P) and Ito *et al.*[3] (Bem1: K482A; Cdc24: D820A and D824A) are indicated in bold. The triangles indicate residues that are part of the interface in the complex and establish intermolecular hydrogen bonds. The green arrow marks a conserved residue with 0% surface-accessibility. The secondary elements are presented on top with orange ovals for α-helices and red rectangles for β-sheets.

that such simulations do not result in any significant improvement, for low levels of sequence homology, as was the case here, MD simulations proved to be a powerful tool to improve the 3D model.

Docking of Bem1 PB1 and Cdc24 PB1

Titration experiments monitored by NMR allowed the identification of residues involved in the binding of the two PB1 domains. In both PB1
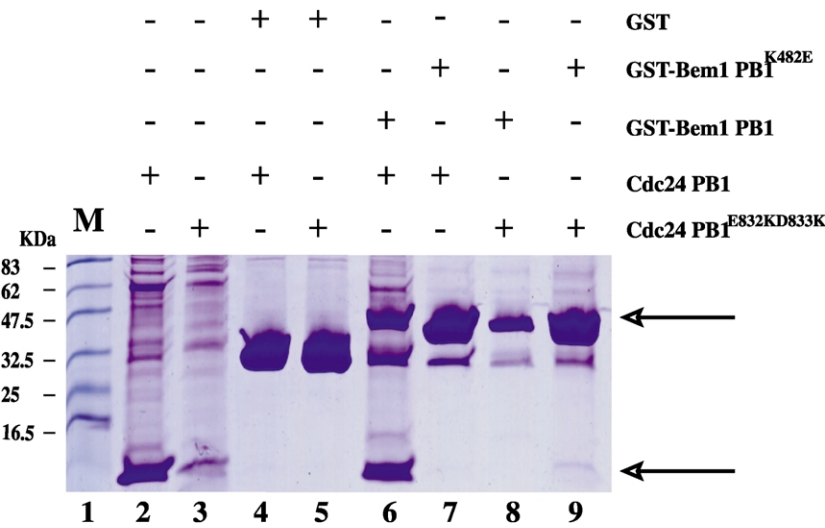

Figure 6. *In vitro* binding assays. GST (lanes 4,5), GST-Bem1 PB1 (lanes 6,8) or GST-Bem1 PB1 K482E (lanes 7,9) were immobilized on glutathione-Sepharose beads, and incubated as indicated with wild-type Cdc24 PB1 (lanes 4, 6 and 7) or Cdc24 PB1 E832K D833K (lanes 5, 8 and 9). After washing, bound proteins were eluted with 3× Laemmli buffer and analyzed by SDS-PAGE and staining with Coomassie brilliant blue. M, protein marker. Note that swapping the charged residues involved in the binding interface partially restores binding of the two PB1 domains.
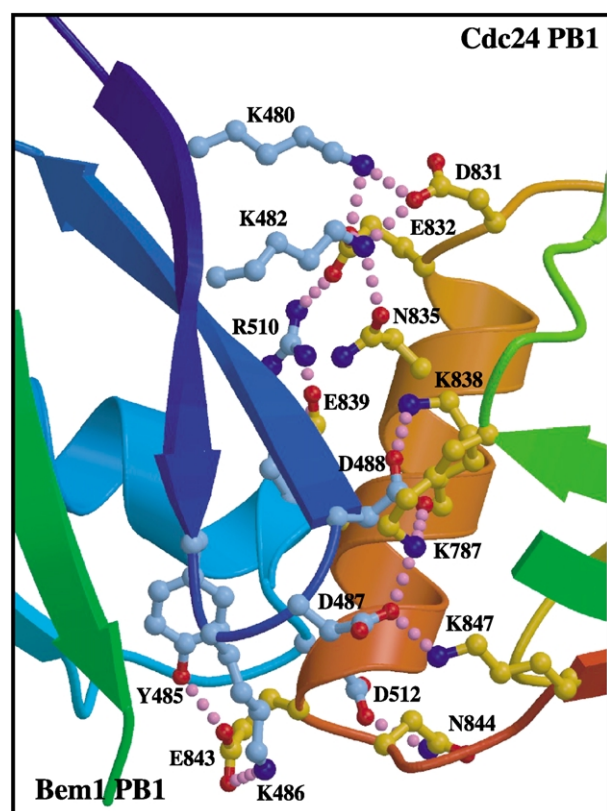
Figure 7. Detailed view of the interface of the Bem1 PB1–Cdc24 PB1 complex. Intermolecular hydrogen bonds and salt-bridges are represented by magenta dotted lines. The orientation of the complex is the same as the top view in Figure 5. This Figure was generated with the programs MOLSCRIPT[37] and Raster3D.[38]

domains, the identified residues cluster on well-defined surfaces of the proteins. Our results are, to a large part, in agreement with those published by Terasawa *et al.*[8] Some noticeable differences are observed, however, for Cdc24 PB1: no significant chemical shift change was detected for residues D820 or D824, which were considered to be essential for the interaction with Bem1 PB1.[8] This apparent discrepancy could be due to the difference in pH conditions between the two studies: Terasawa *et al.*[8] worked at pH 6.3, while the present work was performed at pH 7.4. In our model, a histidine residue (H806) is found in the proximity of the interface. The two aspartate residues, D820 and D824, for which, contrary to the previous study,[8] we did not observe any significant chemical shift perturbations, are lining the interface between Bem1 and Cdc24. They are, however, not involved in any direct intermolecular contacts. We have evidence that they might be involved in the interface of a Cdc24 PB1 homodimer. Histidine 806 could thus govern the pH-dependent homodimerization of Cdc24 PB1 (A.v.D.-P. *et al.*, unpublished results).

HADDOCK: a powerful method for predicting the structure of protein complexes

Traditional *in silico* docking approaches rely on defining the interface of complexes based on the surface geometry complementarity and amino acid pairwise affinities of the 3D structure of the unbound molecules[21,22] without *a priori* inclusion of any experimental information. If used, experimental information is typically included only *a posteriori* to filter solutions. HADDOCK, however, is a new method that allows directly incorporation of biological and/or biophysical information to drive the docking.[13] In our case, the ambiguous interaction restraints were derived from NMR titration data. The initial results were based on a 19% identity homology model of Cdc24. These were found to be in very good agreement with results from a later docking with the recently published NMR structure of Cdc24. Indeed, the binding mode, the respective interfaces of the two proteins and about 50% of the intermolecular contacts were identical between the homology-based and the NMR-based solutions. This indicates clearly that HADDOCK is a rather robust docking approach that can accommodate imperfections in the structure and still produce reasonable solutions. Our docking results provide an overview of the complex and a more precise view of the interface between Bem1 PB1 and Cdc24 PB1. In our model, the binding interface consists of complementary charged surfaces with a network of hydrogen bonds and salt-bridges. Among those, two residues are particularly noticeable, $K480^{Bem1\ PB1}$ and $K482^{Bem1\ PB1}$, which interact with $D831^{Cdc24\ PB1}$ and $E832^{Cdc24\ PB1}$. These two lysine residues have been shown to be key residues, since mutations of K480E and K482A abolish the complex formation.[3,12] Our mutagenesis and swapping experiments were based on contact information from the homology-based docking solution. Because of that, $E832^{Cdc24\ PB1}$ and $D833^{Cdc24\ PB1}$ were selected. The latter residue, however, is no longer involved in intermolecular contacts in the NMR-based model. Despite this, and because $E832^{Cdc24\ PB1}$ is a key residue, our mutagenesis experiments were successful and swapping those charged residues in Bem1 PB1 and Cdc24 PB1 could restore, to some extent, binding of the two domains, providing support for our model. If the NMR structure of Cdc24 had been available when this work was initiated, clearly $D831^{Cdc24\ PB1}$ would have been selected for mutagenesis instead of $D833^{Cdc24\ PB1}$ and the swapping experiment would have involved $K480^{Bem1\ PB1}$. This would most likely have resulted in a stronger restoration of binding than what was observed with the chosen mutants.

PB1–PB1 interaction domains

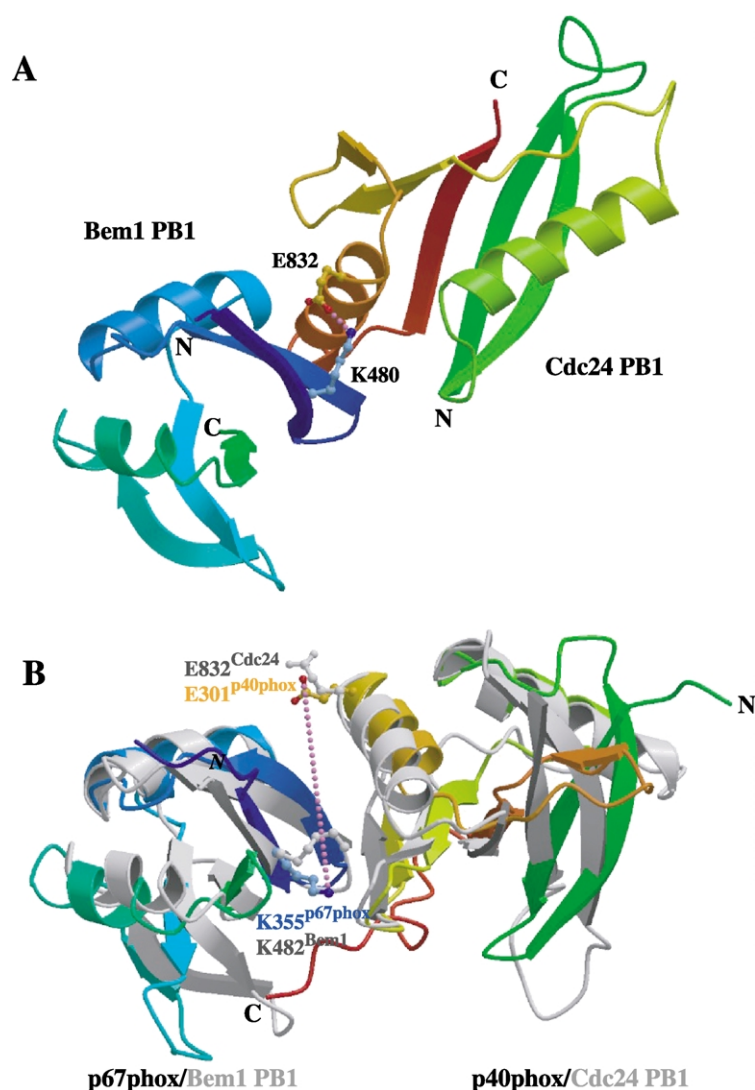Sequence alignments revealed the presence of highly conserved regions embedded in PB1

Figure 8. Comparison of the Bem1−Cdc24 PB1 domains complex with the p67phox−p40phox PB1 crystal structure[10] (PDB entry 1OEY). A, Model of Bem1-Cdc24 PB1 domains complex obtained with HADDOCK using NMR titration data. B, Crystal structure of the p67phox−p40phox PB1 domains complex in color with superimposed Bem1 and Cdc24 in gray. Bem1 PB1 and Cdc24 PB1 were superimposed separately on p67phox and p40phox, respectively. The orientation of p67phox/Bem1 is the same in the two parts of the Figure. K482[Bem1] and E832[Cdc24] are connected by a magenta dotted line. These residues have been shown from mutagenesis data to interact (see the text). The corresponding residues in p67phox, K355, and in p40phox, E301, are more than 11 Å apart. This Figure was generated with the programs MOLSCRIPT[37] and Raster3D.[38]

domains, termed OPR and OPR-binding motifs.[18] Mutagenesis experiments followed by *in vitro* binding assays, docking approaches, and X-ray crystallography studies highlighted their crucial roles in PB1 domain association.[10,11] One characteristic of these two conserved motifs is that they contain charged residues. Indeed, the OPR motif carries acidic residues (aspartate and glutamate), whereas the OPR-binding domain contains mainly basic residues (lysine and arginine). These charged residues enable the establishment of electrostatic interactions between both conserved regions, leading to complex formation.[11] In our model of the Bem1−Cdc24 PB1 complex, all residues playing a role at the interface are located in the highly conserved OPR and OPR-binding motif of Cdc24 PB1 and Bem1 PB1 domains, respectively.

Since PB1 domains carry highly conserved motifs, we address the question of whether residues involved in heterodimerization are conserved. Analysis of the recent crystal structure of p67phox/p40phox PB1 complex reveals that four lysine residues stabilize the heterodimer,

K355[p67phox PB1], K382[p67phox PB1], K365[p67phox PB1] and K418[p67phox PB1], by establishing electrostatic interactions with the p40phox OPR motif. However, only one of these four crucial lysine residues is conserved in Bem1 PB1 (corresponding to K482[Bem1 PB1]). Our model of the Bem1−Cdc24 PB1 complex can be compared to the p67phox/ p40phox PB1 crystal structure to address the question of the conservation of the binding mode between PB1 domains (Figure 8). Although, in both cases, acidic and basic clusters are present at the interface they differ significantly. For example, in the Bem1−Cdc24 PB1 complex, the second acidic cluster (D831−D833) of Cdc24 PB1 plays an essential role in establishing electrostatic interactions with Bem1 PB1, whereas its first acidic cluster (D820−D824) is silent. In contrast, in p40phox−p67phox PB1, the first acidic cluster of p40phox (D289−D293) is of major importance. As a consequence of these differences, we were not able to observe a conservation of contact at the residue level. Interestingly, the sequence alignment based on the superimposition of the Cdc24 PB1

(757–854) NMR structure onto p40phox PB1 crystal structure results in a shift of 11 residues compared to the usual sequence alignment found in the literature. On the basis of the structure alignment, it is likely that the second acidic cluster of p40phox PB1 corresponds to the first one of Cdc24 PB1. This observation again supports the idea of a different binding mode between these PB1 domain complexes. Particularly noteworthy in the Cdc24–Bem1 PB1 complex is the role of the K480-K482$^{Bem1\ PB1}$–D831-E832$^{Cdc24\ PB1}$ interaction in the heterodimer stability, as shown in Figure 7. The structure-based sequence alignment reveals that K482$^{Bem1\ PB1}$ corresponds to K355^{p67phox} and E832$^{Cdc24\ PB1}$ to E301^{p40phox}. Although these residues (K355^{p67phox} and E301^{p40phox}) are located at the interface of the p67phox–p40phox PB1 complex, they are more than 11 Å distant from each other, even when considering the most favourable side-chain conformations (Figure 8B). In a model of the Bem1–Cdc24 PB1 domain complex built onto the p67phox–p40phox PB1 complex, only few of the residues identified from NMR titration experiments to be involved in the binding would make intermolecular contacts. In addition, such a binding mode would involve residues that are not shifting in our NMR titration experiments such as D487 and K545 for Bem1 PB1, and PB1 K815, D822 and D824 for Cdc24 (the latter two correspond to the first acidic cluster). A binding mode of Bem1 and Cdc24 similar to the p67phox–p40phox PB1 complex is therefore most unlikely and would be inconsistent with our mutagenesis data.

Conservation at the residue level would imply that PB1 domains could be interchangeable, meaning that the Bem1 PB1 domain could be a substitute for the p67phox PB1 domain, for example. This hypothesis is not valid, however, as has been confirmed by Ito *et al.*[3] Indeed, Bem1 PB1 interacts specifically with Cdc24 PB1 but does not bind to either ZIP PB1 domain, Scd1 PB1 or p40phox PB1. Hydrophobic residues often mediate interactions conferring specificity. In our model of the Bem1–Cdc24 PB1 complex, L841$^{Cdc24\ PB1}$ and Y485$^{Bem1\ PB1}$ provide the major hydrophobic interaction. In addition, both A491 and R510$^{Bem1\ PB1}$, which have been proposed to account for the specificity of this interaction,[23] are making intermolecular contacts. In the case of the p67phox/p40phox PB1 complex, specificity is mainly due to the presence at the interface of R397^{p67phox}/P339^{p40phox} and W425^{p67phox}/Y337^{p40phox}.[10] It is therefore tempting to conclude that complex formation of PB1 domains is conserved at the motif level (OPR and OPR-binding motif) but not necessarily at the residue level. This observation has implications for specificity. All PB1 pairs possibly present specific molecular interactions that make them unique. This makes a generalization difficult and additional structures of PB1 domains complexes will be required to unravel the issue of binding mode and specificity.

Materials and Methods

Molecular biology techniques

All mutations were performed by PCR-based site-directed mutagenesis and confirmed by sequencing. The K482E mutation was introduced into the pGEX-4–T-Bem1 PB1 plasmid (pNP141), whereas mutations E832K-D833K of Cdc24 PB1 were obtained using the plasmid pGEX-4-T-Cdc24 PB1 (pNP226). The oligonucleotides were synthesized by Invitrogen (Carlsbad, CA).

Expression and purification of Bem1 PB1 and Cdc24 PB1 domains

Cdc24 PB1 (780–854) and Bem1 PB1 (472–551) were produced as GST fusions in the *E. coli* strain BL21. *E. coli* were grown in LB (Luria-Bertani) medium for the production of unlabeled protein, while commercial minimal medium (Silantes) was used to produce ^{15}N-labeled protein. Protein expression was induced by adding isopropyl-β,D-thiogalactopyranoside (IPTG) to 1 mM at 37° C. After cell lysis, the debris was pelleted by centrifugation at 10,000g for five minutes. The supernatant was then incubated for two hours at 4° C with glutathione-agarose beads (Pharmacia). After several washes with thrombin buffer (50 mM Tris–HCl (pH 7.5), 100 mM NaCl, 0.5 mM CaCl$_2$, 10% (v/v) glycerol), the GST was cleaved at 37° C using ten units of human thrombin (Sigma). Cdc24 PB1 and Bem PB1 were further purified by ion-exchange chromatography. In both cases the proteins were eluted with a gradient of 0.1 M–1.0 M NaCl. The purified proteins were then dialysed against the buffer used for NMR experiments containing 100 mM NaCl, 20 mM sodium phosphate, 1 mM EDTA and 0.01 × proteases inhibitor cocktail (Roche) at pH 7.3. The final sample concentration was 1 mM protein in H$_2$O/^{2}H$_2$O (9:1, v/v).

In vitro GST-binding assays

All proteins were expressed as GST-fusion proteins and the purification of Cdc24 PB1 and Cdc24 PB1E832K-D833K was performed as described below. For GST-Bem1 PB1 and GST-Bem1 PB1K482E, the crude extract was incubated with Glutathione-Sepharose 4B beads (Amersham Biosciences) at 4° C for two hours. The resin was recovered by centrifugation, washed extensively with 150 mM NaCl, 50 mM Tris–HCl (pH 7.5), 1 μM EGTA and a protease inhibitor cocktail (Tablets "complete" from Roche) to remove non-specifically bound proteins. Finally, the GST-fusion proteins were eluted with reduced glutathione (20 mM in Tris–HCl, pH 8). For the GST-binding assays, purified proteins were incubated in solution A (50 mM Tris–HCl (pH 7.5), 150 mM NaCl) for two hours at 4° C. Glutathione-Sepharose 4B beads were added to the mixture and incubated for an additional two hours. After several washes with solution A, bound proteins were eluted with 3 × Laemmli buffer and analyzed by SDS-PAGE.[24]

NMR spectroscopy

Titration NMR experiments were performed using ^{1}H/^{15}N HSQC spectra of Cdc24 PB1 and Bem1 PB1 on a Bruker 600 MHz spectrometer, at 300 K. A matrix of

(64 × 512) complex points in (*t*1, *t*2) was obtained with 16 scans/ free induction decay (FID). A 1 mM stock solution of unlabeled Cdc24 PB1 was added in a stepwise manner to a 1 mM ^{15}N-Bem1 PB1 sample. We processed materials in a similar way for the ^{15}N-Cdc24 PB1/Bem1 PB1 titration. Combined ^{1}H/^{15}N chemical shift perturbations were calculated using the relationship:[25]

$$\Delta\delta = [(\Delta\delta_{1HN})^2 + (\Delta\delta_{15Nppm}/6.51)^2]^{1/2}$$

Combined chemical shift changes ($\Delta\delta$) were considered significant if they exceeded the average chemical shift change plus one standard deviation. The 3D NOESY-HSQC experiment was recorded from a ^{15}N-Cdc24 PB1 sample (1 mM) on a Bruker 750 MHz spectrometer, at 300 K, with 100 ms mixing time. All NMR spectra were processed using the software package NMRpipe[26] and analyzed using NMRview.[27] The ^{N}H, ^{15}N assignment of the Bem1 PB1 and Cdc24 PB1 spectra are based on the assignment published by Terasawa *et al*.[8]

Molecular dynamics simulation

The theoretical model derived from the sequence alignment with the Ras-binding domain of Ral GEF was used as starting point for the molecular dynamics simulations in explicit solvent. The simulations were run with the GROMACS 3.1 MD package[28] using the GROMOS96 force-field.[29] The structure was solvated in a cubic box of SPC water[30] using a minimum distance of 1.4 nm between the protein and the box edges.

After a first steepest descent energy minimization with positional restraints on the solute, six sodium counterions (Na$^+$) were introduced to obtain an electro-neutralized system. A second energy minimization was performed, followed by five successive 20 ps MD equilibration runs. During these, the position restraints force constant on the solutes, K_{posre}, were decreased progressively (1000, 1000, 100, 10, 0 kJ mol^{-1} nm^{-2}). A 10 ns production run was then performed at constant temperature (300 K) and pressure (1 atm) with weak coupling (0.1 and 1 ps^{-1}) to reference T and P baths[31] using a 4 fs time-step for the integration of the equations of motion. Non-bonded interactions were calculated using twin range cutoffs of 0.8 nm and 1.4 nm. Long-range electrostatic interactions beyond the cutoff were treated with the generalized reaction field model using a dielectric constant of 54.[32] The LINCS algorithm[33] was used for bond length constraint, in conjunction with dummy atoms for the aromatic rings and amino groups in side-chains, enabling the use of the 4 fs integration time-step.[34] After 10 ns, the observed NOEs were introduced as distance restraints and the simulation was extended for 5 ns to reach a total of 15 ns.

Docking

Structural coordinates

Bem1 PB1 coordinates were taken from the RCSB Protein Data Bank, entry 1ipg containing 20 models. The first ten flexible residues were removed for docking. For the Cdc24 PB1 domain, ten conformations, taken every 0.5 ns for the last 5 ns of the MD simulation (10.5–15 ns), were used as starting point for docking.

Docking protocol

The docking was performed with HADDOCK 1.2, supporting docking from ensembles of structures†. The active residues used to define the ambiguous interaction restraints (AIRs) were those showing significant ^{1}H/^{15}N chemical shift perturbations in the HSQC titration experiments and solvent-accessible (see Results). The average relative residue solvent-accessible surface area values (*avRSA*) were calculated with NACCESS‡ over the ensemble of structures. Only the residues having either backbone or side-chain *avRSA* + standard deviation above 50% were retained for the definition of active and passive residues. The passive residues correspond to the residues that are surface neighbors of the active residues as identified by visual inspection and that have a high level of solvent-accessibility ($>$50%). The interface amino acid residues, which constitute the flexible segments, are defined by the active and passive amino acid residues used in the AIRs ± 2 sequential residues.

Since the docking was performed from 20 and ten conformations of Bem1 PB1 and Cdc24 PB1, respectively, 1600 rigid-body docking solutions were generated (each of the 200 possible combinations of starting conformations was thus used eight times). In addition, for each of the starting conformations, ten rigid-body trials were performed and only the best solution based on the intermolecular energy was kept, bringing the total effective docking trials to 16,000. The best 200 solutions, sorted according to the intermolecular energy (sum of van der Waals, electrostatic and AIRs energy terms), were further submitted to the semi-flexible simulated annealing and water refinement as described.[13] The solutions were clustered using a 1.5 Å RMSD cut-off criterion and the ten lowest-energy structures of the lowest intermolecular energy cluster were selected for analysis. Intermolecular contacts (hydrogen bonds and non-bonded contacts) were analyzed with DIMPLOT, which is part of the LIGPLOT software,[35] using the default settings (3.9 Å heavy-atoms distance cut-off for non-bonded contacts; 2.7 Å and 3.3.5 Å proton–acceptor and donor–acceptor distance cut-offs respectively with minimum 90° angles (D–H–A, H–A–AA, D–A–AA) for hydrogen bonds).[36]

† http://www.nmr.chem.uu.nl/haddock
‡ http://wolf.bms.umist.ac.uk/naccess/

References

1. Pawson, T., Raina, M. & Nash, P. (2002). Interaction domains: from simple binding events to complex cellular behavior. *FEBS Letters*, **513**, 2–10.
2. Peterson, J., Zheng, Y., Bender, L., Myers, A., Cerione, R. & Bender, A. (1994). Interactions between the bud emergence proteins Bem1p and Bem2p and Rho-type GTPases in yeast. *J. Cell Biol.* **127**, 1395–1406.
3. Ito, T., Matsui, Y., Ago, T., Ota, K. & Sumimoto, H. (2001). Novel modular domain PB1 recognizes PC motif to mediate functional protein–protein interactions. *EMBO J.* **20**, 3938–3946.
4. Pruyne, D. & Bretscher, A. (2000). Polarization of cell growth in yeast. *J. Cell Sci.* **113**, 571–585.
5. Nakamura, R., Sumimoto, H., Mizuki, K., Hata, K., Ago, T., Kitajima, S. *et al.* (1998). The PC motif: a novel and evolutionarily conserved sequence involved in interaction between p40phox and p67phox, SH3 domain-containing cytosolic factors of the phagocyte NADPH oxidase. *Eur. J. Biochem.* **251**, 583–589.
6. Gong, J., Xu, J., Bezanilla, M., van Huizen, R., Derin, R. & Li, M. (1999). Differential stimulation of PKC phosphorylation of potassium channels by ZIP1 and ZIP2. *Science*, **285**, 1565–1569.
7. Sanz, L., Sanchez, P., Lallena, M. J., Diaz-Meco, M. T. & Moscat, J. (1999). The interaction of p62 with RIP links the atypical PKCs to NF-κB activation. *EMBO J.* **18**, 3044–3053.
8. Terasawa, H., Noda, Y., Ito, T., Hatanaka, H., Ichikawa, S. *et al.* (2001). Structure and ligand recognition of the PB1 domain: a novel protein module binding to the PC motif. *EMBO J.* **20**, 3947–3956.
9. Ponting, C. P., Ito, T., Moscat, J., Diaz-Meco, M. T., Inagaki, F. & Sumimoto, H. (2002). OPR, PC and AID: all in the PB1 family. *Trends Biochem. Sci.* **27**, 10.
10. Wilson, M. I., Gill, D. J., Perisic, O., Quinn, M. T. & Williams, R. (2003). PB1 domain-mediated heterodimerization in NADPH oxidase and signaling complexes of atypical protein kinase C with Par6 and p62. *Mol. Cell*, **12**, 39–50.
11. Lamark, T., Perander, M., Outzen, H., Kristiansen, K., Overvatn, A., Michaelsen, E. *et al.* (2003). Interaction codes within the family of mammalian Phox and Bem1p domain-containing proteins. *J. Biol. Chem.* **278**, 34568–34581.
12. Butty, A. C., Perrinjaquet, N., Petit, A., Jaquenoud, M., Segall, J. E., Hofmann, K. *et al.* (2002). A positive feedback loop stabilizes the guanine-nucleotide exchange factor Cdc24 at sites of polarization. *EMBO J.* **21**, 1565–1576.
13. Dominguez, C., Boelens, R. & Bonvin, A. M. (2003). HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737.
14. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data Bank. *Nucl. Acids Res.* **28**, 235–242.
15. Xu, Y. & Xu, D. (2000). Protein threading using PROSPECT: design and evaluation. *Proteins: Struct. Funct. Genet.* **40**, 343–354.
16. Sali, A. & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
17. Yoshinaga, S., Kohjima, M., Ogura, K., Yokochi, M., Takeya, R., Ito, T. *et al.* (2003). The PB1 domain and the PC motif-containing region are structurally similar protein binding modules. *EMBO J.* **22**, 4888–4897.
18. Ponting, C. P. (1996). Novel domains in NADPH oxidase subunits, sorting nexins, and PtdIns 3-kinases: binding partners of SH3 domains? *Protein Sci.* **5**, 2353–2357.
19. Chant, J. (1994). Cell polarity in yeast. *Trends Genet.* **10**, 328–333.
20. Gulli, M. P., Jaquenoud, M., Shimada, Y., Niederhauser, G., Wiget, P. & Peter, M. (2000). Phosphorylation of the Cdc42 exchange factor Cdc24 by the PAK-like kinase Cla4 may regulate polarized growth in yeast. *Mol. Cell*, **6**, 1155–1167.
21. Halperin, I., Ma, B., Wolfson, H. & Nussinov, R. (2002). Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins: Struct. Funct. Genet.* **47**, 409–443.
22. Janin, J., Henrick, K., Moult, J., Ten Eyck, T., Sternberg, M. J. E., Vajda, S. *et al.* (2003). CAPRI: a critical assessment of predicted interactions. *Proteins: Struct. Funct. Genet.* **52**, 2–9.
23. Noda, Y., Kohjima, M., Izaki, T., Ota, K., Yoshinaga, S., Inagaki, F. & Ito, T. (2003). Molecular recognition in dimerization between PB1 domains. *J. Biol. Chem.* **278**, 43516–43524.
24. Harlow, E. & Lane, D. (1988). *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
25. Mulder, F. A. A., Schipper, D., Bott, R. & Boelens, R. (1999). Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilisins*. *J. Mol. Biol.* **292**, 111–123.
26. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
27. Johnson, B. A. & Blevins, R. A. (1994). NMRview: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
28. Lindahl, E., Hess, B. & van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**, 306–317.
29. Daura, X., Mark, A. E. & van Gunsteren, W. F. (1998). Parameterization of aliphatic CHn united atoms of GROMOS96 force field. *J. Comput. Chem.* **19**, 535–547.
30. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & Hermans, J. (1981). Interaction models for water in relation to protein hydration. In *Intermolecular Forces, B* (Pullman, ed.), pp. 331–342, Reidel, Dordrecht.
31. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
32. Tironi, I. G., Sperb, R., Smith, P. E. & van Gunsteren, W. F. (1995). A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459.
33. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472.
34. Feenstra, K. A., Hess, B. & Berendsen, H. J. C. (1999). Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **20**, 786–798.

35. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
36. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127–134.
37. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
38. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.