Functional specificity conferred by the unique plasticity of fully α-helical Ras and Rho GAPs

Michel Souchet[a], Anne Poupon[b], Isabelle Callebaut[c], Isabelle Léger[a], Jean-Paul Mornon[c,*], Antoine Bril[a], Thierry P.G. Calmels[a,1]

[a] *SmithKline Beecham Laboratoires Pharmaceutiques, 4 rue Chesnay-Beauregard, 35760 Saint-Grégoire, France*
[b] *Département d'Ingénierie et d'Etude des Protéines, Bat. 152, CEA/Saclay, 91191 Gif sur Yvette Cedex, France*
[c] *Systèmes moléculaires et Biologie structurale, LMCP, CNRS UMR C7590, UP6, Case 115, 4 Place Jussieu, 75252 Paris Cedex 05, France*

Received 6 June 2000

Edited by Pierre Jolles

Abstract Structural comparisons of the two GTPase activating proteins (GAPs) p120 and p50 in complex with Ras and Rho, respectively, allowed us to decipher the functional role of specific structural features, such as helix α8c of p120 and helix A1 of p50, necessary for small GTPase recognition. We identified important residues that may be critical for stabilization of the GAP/GTPase binary complexes. Detection of topohydrophobic positions (positions which are most often occupied by hydrophobic amino acids within a family of protein domains) conserved between the two GAP families led to the characterization of a common flexible four-helix bundle. Altogether, these data are consistent with a rearrangement of several helices around a common core, which strongly supports the assumption that p50 and p120 GAPs derive from a unique fold. Considered as a whole, the remarkable plasticity of GAPs appears to be a means used by nature to accurately confer functional specificity. **© 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.**

Key words: Four-helix bundle; Plasticity; GTPase activating protein; Ras; Rho

1. Introduction

Small G proteins of the Ras superfamily control a wide range of biological functions and operate as molecular switches which cycle between GTP-bound and GDP-bound states. GDP/GTP exchange is promoted by guanine nucleotide exchange factors that allow GTPases to interact with effector molecules. Hydrolysis of bound GTP is intrinsically a very slow process which is accelerated by several orders of magnitude through interaction with GTPase activating proteins (GAPs). Until recently, such modulators of small GTPases were not considered to be structurally related, although they act on partners like Ras or Rho which display high structural similarity. The structures of several GAPs have been characterized, including those of p120 [1] and p50 [2] which are specific for the Ras and Rho subfamily, respectively. Their interacting domains in contact with small G proteins carry a crucial catalytic arginine [3]. These domains have been shown to be fully α-helical, as illustrated by X-ray analysis of apo or complexed p120 (specific for Ras) and p50 (specific for Rho) GAPs, revealing modules of nine and eight helices, respectively [1,2,4,5]. Furthermore, we and others have reported that p120 and p50 GAPs are related, supporting a shared ancestry [6–8]. In the present study, we further describe the common structural relationship of the two GAPs at the sequence and three-dimensional (3D) structural levels and characterize key structural elements involved in a rare malleability of α-helical packings. It is noteworthy that the overall plasticity of GAP-interacting domains, which spreads over a rather large α-helical domain, may play a key role in the specific recognition of the GTPase partner by the respective GAP.

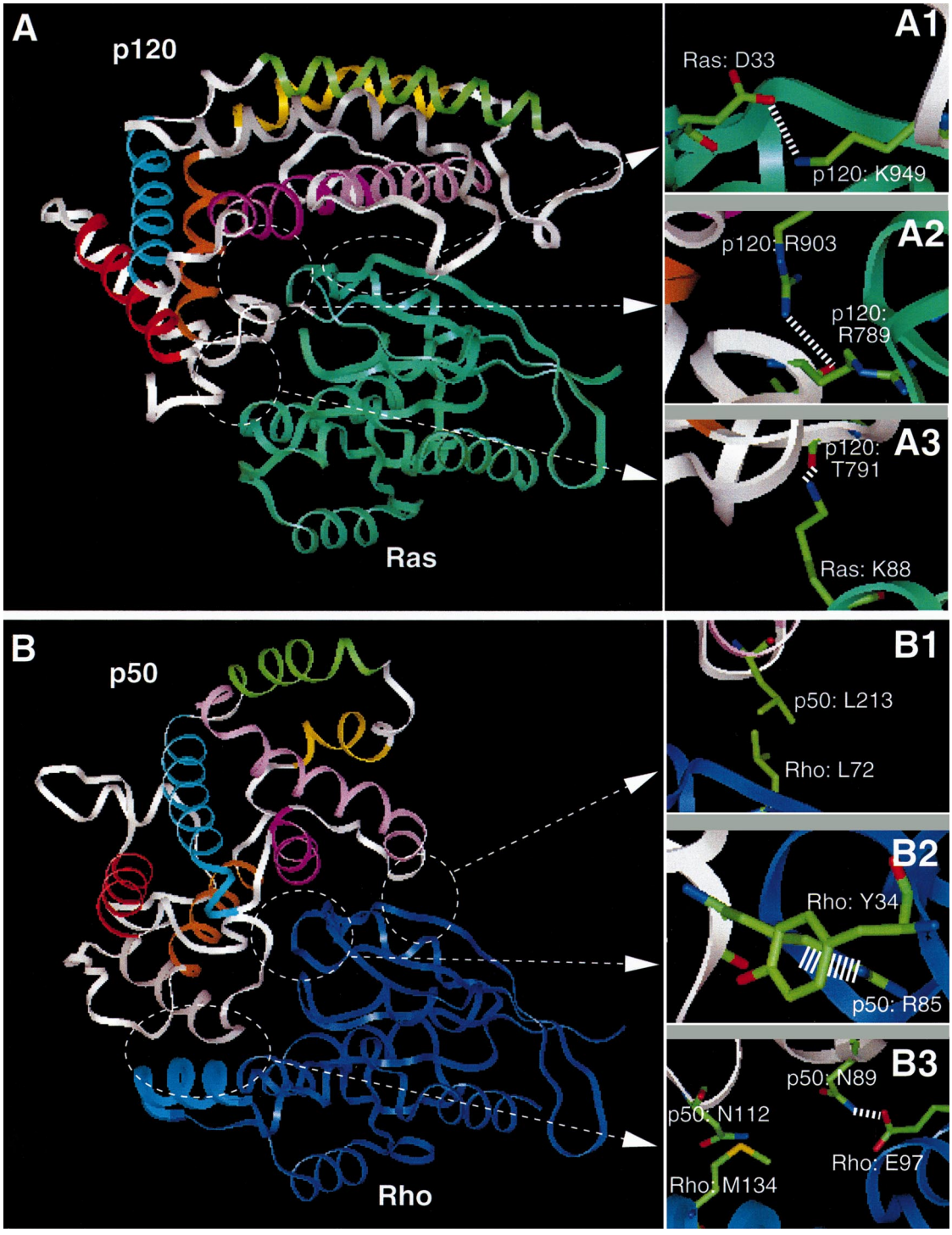
2. Materials and methods

The sequence alignment of p120Ras and p50Rho GAP-interacting domains was performed by hydrophobic cluster analysis (HCA) for which the methodology is reviewed in [9] and exemplified by new studies [10–12]. This sensitive bidimensional alignment method led to the detection of coherent sequence similarities between the interacting domains of both GAPs at a level currently observed for distantly related proteins: 15% sequence identity, 65% hydrophobic matching and statistical *Z*-scores of 4.2, 5.4 and 6.0 σ for identity, similarity (Blosum 62 matrix) and hydrophobic matching, respectively. The resulting sequence reliability index of 4.9 represents the ratio between the observed product of the three *Z*-scores and the best equivalent product for 10 000 random shuffled sequences. Fig. 3A was carried out with the ESPript package [13]. The crystal structures used in this work have the following PDB entry codes: 1WER for p120RasGAP, 1WQ1 for co-crystal p120RasGAP/Ras/GDP, 1RGP for p50RhoGAP, 1TX4 for co-crystal p50RhoGAP/Rho/GDP, 1NF1 for the GAP-related domain of neurofibromin (NF-1), 1PBW for the P85α subunit BCR-homology domain. p120 (1WQ1), neurofibromin (1NF1) Ras GAPs and p50 (TX4), BCR (1PBW) Rho GAPs were used to identify the hydrophobic residues in topohydrophobic positions in each subfamily (see legend of Fig. 2 for the definition of topohydrophobic positions and associated references). Distance, dispersion and accessible surface calculations for topohydrophobic positions were performed using the program CHAP [14]. Solvent accessibilities were computed by the method of Lee and Richards [15]. The rendering displayed in Figs. 1 and 3B as well as distance measurements performed with 3D structures were performed with the InsightII 980 package (Molecular Simulations Inc., San Diego, CA, USA). The four-helix bundle is defined according to previous definitions [16]. The so-called cradle fold (three-helix core) is defined as the invariant and common scaffold of p120 α2c-α3c-α4c/p50 B-C-D helices [7].

*Corresponding author. Fax: (33)-1-44 27 37 85.
E-mail: mornon@lmcp.jussieu.fr

[1] Also corresponding author. Fax: (33)-2-99 28 04 44;
E-mail: thierry_calmels@sbphrd.com

0014-5793/00/$20.00 © 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.
PII: S0014-5793(00)01769-5



3. Results

3.1. Structural comparison of the two GAPs in complex with their respective GTPase

In order to determine common features between the two GAPs, a structure comparison was performed with the two binary complexes p120/Ras and p50/Rho (Fig. 1A,B) using Ras and Rho GTPases as reference (Fig. 1C). (i) The two arginine fingers are oriented very similarly as shown in Fig. 1C2 upon superimposition of the two binary complexes (Fig. 1C). Interestingly, such a topological similarity around the crucial arginines was also observed for p120 and p50 in the absence of the GTPase partner [7]. Therefore, these catalytic fingers appear as essential pivots for the two complexes with respect to their overall flexibility. (ii) Although the two GAP structures have marked differences, the four-helix bundles superimpose fairly well (rmsd = 4.6 Å) as shown as red, orange, light blue and purple helices in Fig. 1C. (iii) The two

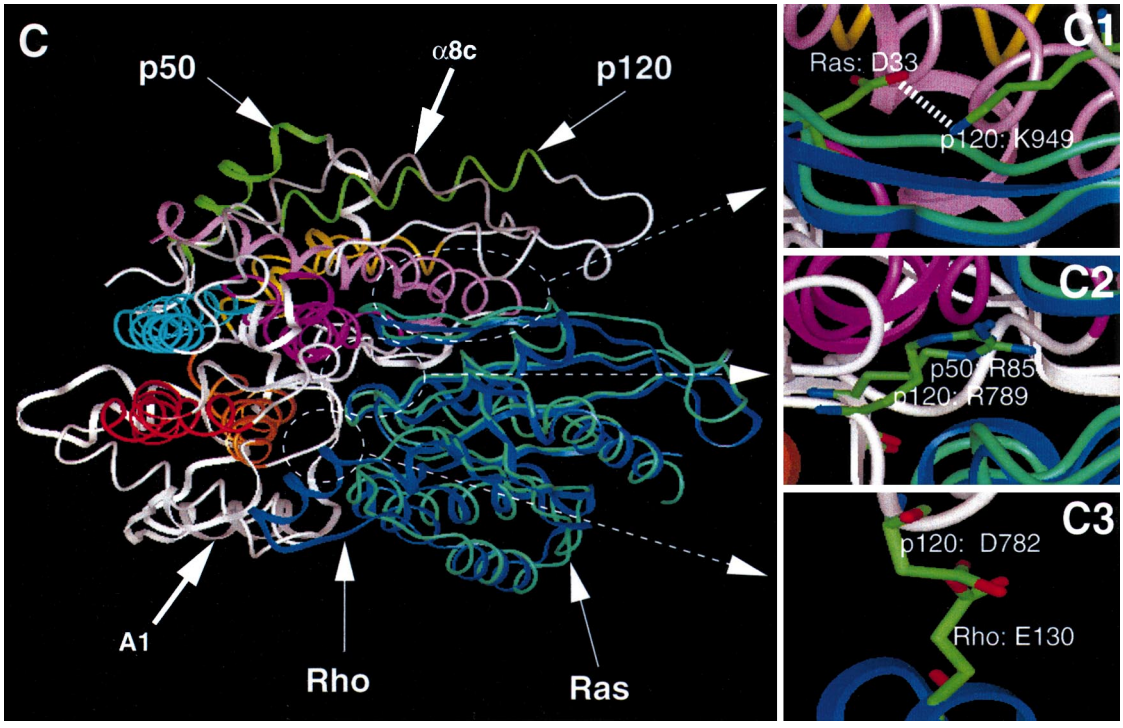
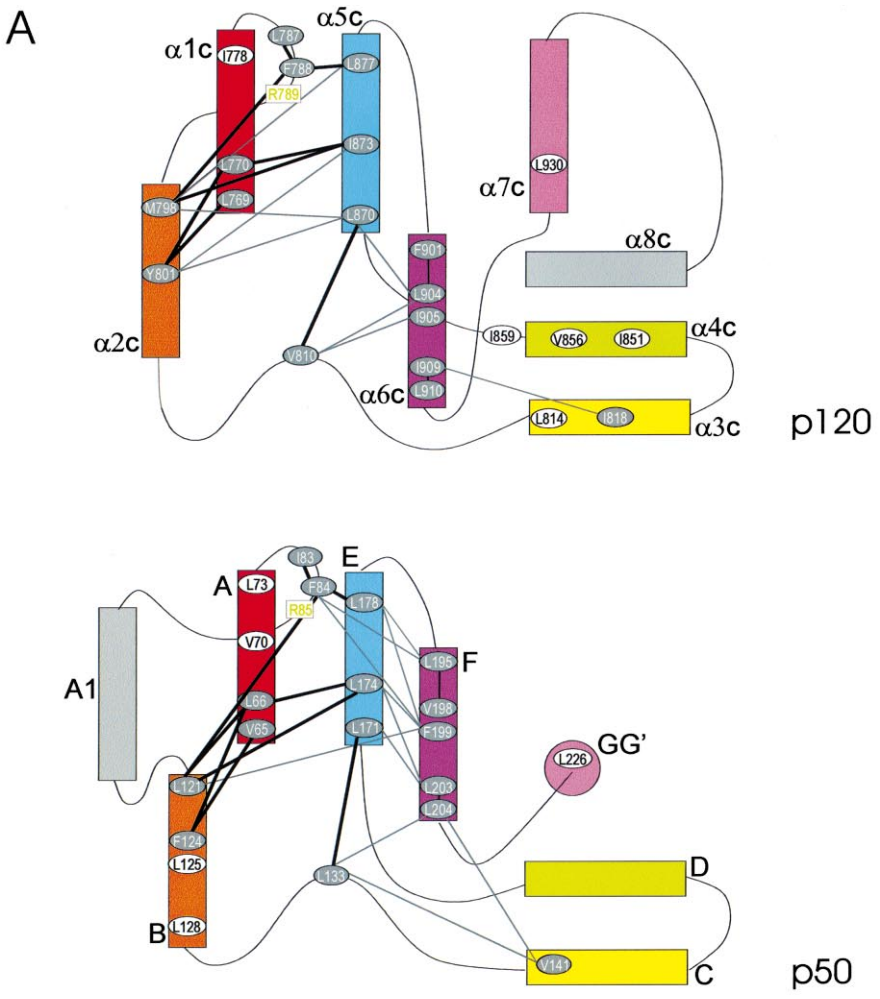


Fig. 1. Comparison of the two X-ray binary complexes p50/Rho and p120/Ras. Important residues are represented with a stick rendering (nitrogen in dark blue, oxygen in red, carbon in green). p120 and p50 are displayed as colored ribbons and identical colors highlight equivalent helices. A: The binary complex of p120 and Ras proteins is displayed as colored ribbons. Orientation of Ras (light green ribbon) is identical to that of Rho (displayed in B). Two stabilizing domains and the area surrounding the catalytic arginine are magnified. A1: The putative salt bridge occurring between p120K949 and RasD33 is indicated by a broken white line. A2: The side chain of p120R903 is shown in hydrogen bond interaction (broken white line) with the carbonyl oxygen atom of the catalytic p120R789. A3: A putative stabilizing hydrogen bond (broken white line) can be identified between p120T791 and RasK88. B: The p50/Rho binary complex is displayed as colored ribbons. Rho is in dark blue except the protruding helix (E125-K135) in medium blue. Three regions at the protein–protein interface, involving either the catalytic arginine or stabilizing areas, are magnified. B1: Putative stabilizing domain involving a hydrophobic contact between p50L213 and RhoL72. B2: The putative π-type interaction between the catalytic arginine p50R85 and RhoY34 is displayed as short white parallel lines. B3: The putative hydrogen bond connecting p50N89 and RhoE97 is indicated by a broken white line (distance cutoff of 3 Å). An additional stabilizing interaction might occur between p50N112 and RhoM134. C: Superimposition of the two X-ray binary complexes, p50/Rho and p120/Ras. The two binary complexes presented in A and B were superimposed via their respective GTPase. The complex p50/Rho is displayed as a solid rectangular ribbon and the complex p120/Ras as a solid circular ribbon. C1: The putative salt bridge occurring between p120K949 and RasD33 appears to be specific to the two protein partners. C2: Superimposition of the two GTPases allows a very good alignment of the two catalytic arginines, p50R85 and p120R789. C3: A steric clash might occur between p120D782 and RhoE130 present in the protruding helix, whenever p120 faces Rho.

complexes reveal that structural specificity of the GAPs is based on the ability to form stabilizing areas which may be distant from the catalytic site. For instance, the protruding extra helix of Rho (segment from E125 to K135) displayed in medium blue in Fig. 1B is in close contact with the loop connecting the extra helix A1 to helix B of p50 (see p50N112/RhoM134; Fig. 1B3) and with the extra helix A1 itself (see p50N89/RhoE97; Fig. 1B3). An additional specific hydrophobic contact is likely to occur between p50L213 and RhoL72 (Fig. 1B1). Similarly, p120K949 located in the loop connecting the extra helix α8c is tightly linked to RasD33 (Fig. 1A1,C1), and RasK88 faces p120T791 which is located in the protruding loop connecting α1c to α2c (Fig. 1A3). Furthermore, in their respective complexes, RhoY34 and p120R903 may contribute to inter- and intra-molecular stabilization of the catalytic arginines (Fig. 1B2,A2), respectively. (iv) Replacement of p50 by p120 in the p50/Rho complex would probably result in formation of steric clashes as shown in Fig. 1C3. This observation may help to explain the selectivity of such modulators with respect to their small GTPase partners.

3.2. A common hydrophobic core

Topohydrophobic positions have recently been defined as positions always occupied by strong hydrophobic residues (V, I, L, F, M, Y, W) within a protein family [14,17]. Identification of such positions was achieved by using two Rho GAPs (p50, BCR) and two Ras GAPs (p120, NF-1) as well as accurate sequence alignments of both superfamilies [7]. Comparison of these positions in the p50 and p120 subfamilies shows a conserved hydrophobic core mainly constituted by 16 equivalent positions: 16 out of 21 and 16 out of 22 topohydrophobic positions for p50 and p120, respectively (displayed in gray in Fig. 2A). In p50, they are located within helices A, B, C, E, F and in loops connecting A to A1 and B to C, whereas the equivalent positions in p120 are located within helices α1c, α2c, α3c, α5c, α6c and in loops connecting α1c to α2c and α2c to α3c. The remaining hydrophobic residues represent positions specific to the Ras and Rho GAP subfamilies. Interestingly, the associated connections of topohydrophobic positions present in helix F of p50 and in helix α6c of p120 appear to be different and specific to each protein subfamily. The intra-molecular relationships displayed as black or gray



B

	Topohydrophobic positions	Non-topohydrophobic positions
Solvent accessibility	2.9 Å^2 (64 hydrophobic aa)	10.8 Å^2 (207 hydrophobic aa)
Mean distance	13.0 Å (480 distances)	21.9 Å (16,744 distances)
Dispersion (*)	1.75 Å (16 positions)	2.98 Å (92 positions)
Closest neighbors (**)	6.36 Å (64 triplets)	

Fig. 2. A: Representation of the topohydrophobic positions network obtained for specific Rho and Ras GAPs. Rectangles represent helices with a color coding similar to Figs. 1 and 3. Conserved topohydrophobic positions are indicated by gray ovals. Topohydrophobic positions found only in one subfamily are displayed as white ovals. Black lines show conserved hydrophobic interactions in both subfamilies and gray lines highlight contacts that are specific to one subfamily. Catalytic arginines (in green) are boxed. B: Characterization of topohydrophobic positions for p120 and p50 GAPs. Main characteristics of hydrophobic amino acids (V, I, L, F, M, Y, W) in topohydrophobic positions versus non-topohydrophobic ones, for both Rho and Ras GAP subfamilies, as deduced from PDB coordinate entries 1WQ1/1NF1 and 1TX4/1PBW, respectively. In the case of non-topohydrophobic positions, only hydrophobic residues are taken into account. Mean distance represents the mean distance between the gravity centers of side chains of residues occupying two positions of the same type (topohydrophobic or not). (*) Dispersion is the mean distance between the gravity centers of side chains of residues occupying a same position in the different proteins of the family: this value was computed separately in the two subfamilies because of important differences observed between the two types of skeleton. (**) Closest neighbors represent the mean distance between one given topohydrophobic position and the two closest topohydrophobic neighbors (distances measured between the gravity centers of side chains of residues): this value is not computed for non-topohydrophobic positions.

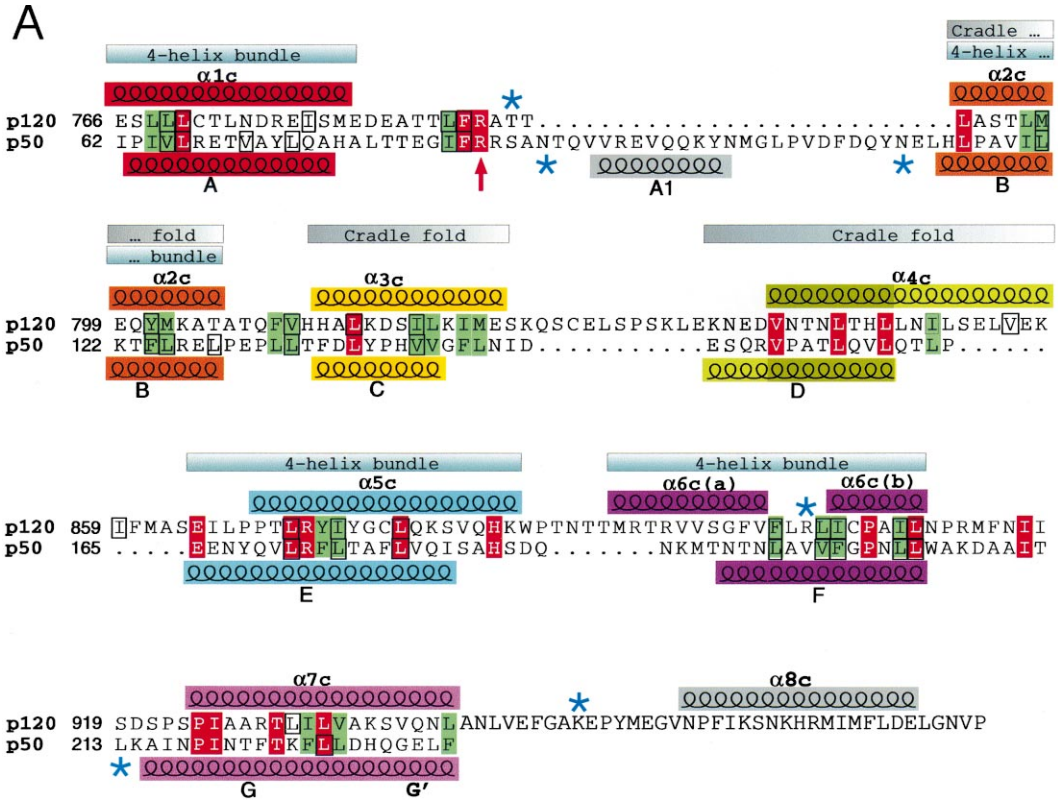


Fig. 3. A: HCA-deduced sequence alignment of p120 and p50 GAP-interacting domains. Identities and hydrophobic conserved positions are shaded in red and green, respectively. The catalytic arginines p120R789 and p50R85 are indicated by a red arrow. Topohydrophobic positions are boxed in each protein sequence. Stars in blue indicate specific stabilizing residues identified in the two structures. α-Helices are delineated above and below the respective sequences (α1c–α8c for p120 and A–G for p50) and a same color is used for the equivalent helices found in the two proteins. Shaded boxes displayed above sequences indicate helices which form the common cradle fold (three-helix core) and/or helices which are part of the flexible four-helix bundle. The darker part of green helix α4c and helix D highlights the co-axial 'sliding effect' observed with VxxxLxxL motif (segment **VPATLQVL** for p50 and **VNTNLTHL** for p120). Extra helices are colored in gray (i.e. A1 and α8c for p50 and p120, respectively). B: Overview of the structural relationship between p50 and p120. The putative rotation/distortion of p50 (top) leads to the p120 topology (bottom). Movement around helices B–C of p50 is indicated by four colored arrows: the tilt of helix α6c (arrow 1 in purple) pushes helix E in the α5c position (arrow 2 in light blue), and in a same anti-clockwise movement forces helix A in the α1c position (arrow 3 in red). Helix G partly follows this general rearrangement and locates in the α7c position (arrow 4 in pink). In addition, the two folds share identical connectivities with two extra helices, A1 (in gray) for p50 inserted between A and B, and the C-terminal helix α8c (in gray) for p120.

lines in both proteins (Fig. 2A) reveal a topohydrophobic position network typical of globular domains [14]. Hydrophobic residues located in topohydrophobic positions are clearly more buried than the others: the mean distance separating topohydrophobic positions is significantly smaller: 13 Å compared to 21.9 Å. Side chains of hydrophobic residues are less dispersed when located in topohydrophobic positions: 1.75 Å compared to 2.98 Å (Fig. 2B).

3.3. Structural plasticity of Ras and Rho GAPs

A sensitive method for sequence alignment of p120 and p50 GAP-interacting domains allowed the detection of coherent sequence similarities between the two α-helical domains sharing 15% sequence identity (Fig. 3A; see Section 2 for details of the methodology). It is noteworthy that identical and similar amino acids are located mainly within the seven overlying α-helices of the two GAPs. Fig. 3B gives an overview of the putative rearrangement of helices occurring between p50 and p120 which may be considered as deriving from a unique fold possessing remarkable flexibility. This is suggested by conserved topohydrophobic positions which form a compact network of interacting side chains (Fig. 2A) defining a hydrophobic core within the four-helix bundle of p50 and p120. In this context, the modified α1c-α2c-α5c-α6c bundle of p120 could be related to the canonical four-helix bundle A-B-E-F in p50 via concerted movements of several helices with respect to the nearly invariant positions of helices α2c and B, taken as reference. The tilt of helix α6c (helix F in p50, Fig. 3B) pushes helix E in the α5c position and in a same anti-clockwise movement, forces helix A in the α1c position. Meanwhile, helix G partly follows this general rearrangement and locates in the α7c position. In addition, the two folds share identical connectivities with two extra helices, A1 for p50 inserted between A and B, and the C-terminal helix α8c for p120.

4. Discussion

4.1. Structural plasticity of Ras and Rho GAPs

This study describes structural specificities between p120 and p50 GAPs and their respective GTPases. For instance, inter-subfamily specificity can be partly characterized by the protruding extra helix of Rho (Fig. 1B) which allows putative stabilizing effects towards p50, distant from the catalytic arginine (such a local contact could not occur between p120 and Rho). On the other hand, an association of p50 and Ras

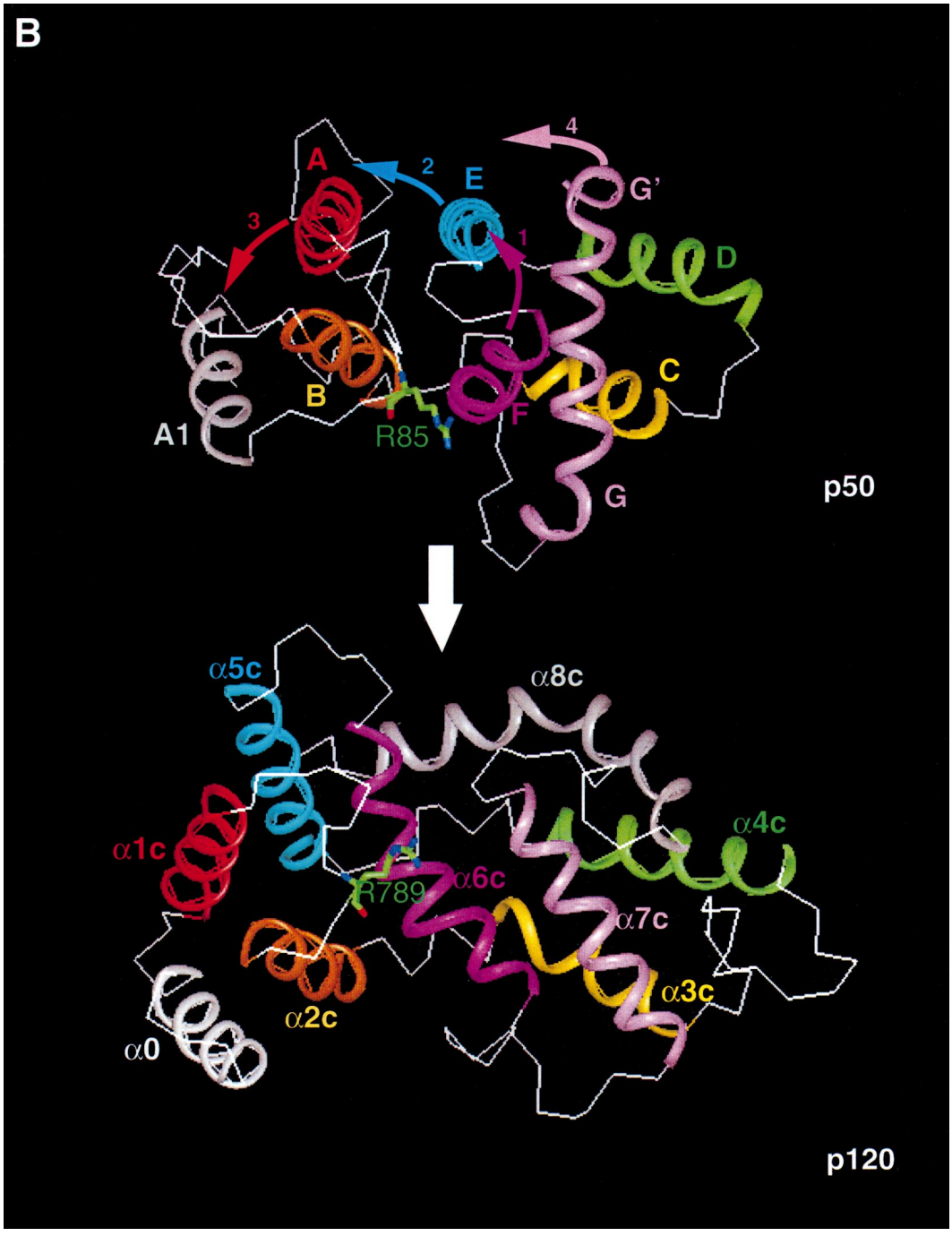


Fig. 3 (*continued*).

would be unfavorable since the stabilizing effect shown for p120/Ras in Fig. 1A1 would probably not take place. The additional contacts displayed in Fig. 1A1, A3, B1 and B3 may participate in molecular recognition of GAPs towards their respective small GTPases and may also be involved in sorting specific partners within each subfamily (intra-subfamily specificity). It may then be assumed that the specificity of the GAP biological function is associated with this plasticity, allowing a similar location of the crucial arginine finger of p120 and p50 (Fig. 1C2) while supplying additional specific stabilizations in the protein–protein interaction process. Furthermore, structure comparisons of p50 and p120 alone or in complex with their respective GTPase led to an rmsd value lower than 0.7 Å for the two GAP-interacting domains. Thus, plasticity is consistent with a stable architecture for p50 and p120 and may be favored by the intrinsic properties of α-helical architecture as reported by Bowie [18]: malleability of α-helix packing may lead to two stable states of a same 'soft' fold and in the present context may drive the Rho/Ras inter-specificity. Moreover, the self-contained network of local H-bonds which forms and characterizes α-helices seems to be more favorable to the winding/unwinding process of coil/α-helix *trans*-conformation than the network of remote H-bonds linking strands within β-sheets. Interestingly, a comparable

unusual plasticity has recently been reported for the quaternary association of α-helices of the Rop (repressor of primer) protein, based on a sole mutation within a loop [19]. Other examples illustrate further such a striking α-helix-associated plasticity linked to functional properties (e.g. movement of the C-terminal helix of retinoic acid nuclear receptors together with variation in its length upon ligand binding [20], unfolding of a helix of the FhuA receptor upon ligand binding [21], large movement of the influenza hemagglutinin associated with the winding/unwinding of secondary structures in the fusion process [22]).

4.2. Sequence versus structure conservation

Due to the exceptional plasticity of the Ras and Rho GAP folds, a discrepancy has been identified between sequence and 3D structure evolutions. This discrepancy, probably due to functional constraints, may be explained by a faster accommodation of 3D structures compared to primary structures. Indeed, the intrinsic malleability of all α-architectures [18] may favor the emergence of considerably different local 3D arrangements although the corresponding sequences conserved clear evolutive relationships. Thus, these two GAP-related domains appear to provide some evidence that sequence is better conserved than 3D structure. This discrepancy is illustrated, for instance, with the 'sliding effect' occurring between helices α4c of p120 and D of p50 (segments in dark green in Fig. 3A). This axial translation of two α-helix turns prevents a direct structural superimposition for the pair of three successive helices composing the so-called cradle fold α2c/B, α3c/C and α4c/D [7].

In conclusion, the present study demonstrates how homofunctionality derives from GAP proteins sharing a flexible scaffold. In this context, specific molecular recognition of the specific small GTPase partner by the respective GAP may result from different packings of helices around a common core, allowing a conserved topology in the vicinity of the crucial catalytic arginine. Thus, the plasticity of the two GAP families represents an important feature of their evolution. This plasticity, possibly triggered by functional constraints, may represent an adaptive mechanism allowing a putative overall dynamic *trans*-conformation or, alternatively, a sole pre-eminence of two states deriving from a common ancestor. In this context, plasticity may be an accurate means for a given GAP in vivo to sort out and recognize partners such as the small GTPases of the Ras and Rho subfamilies, which have highly similar structures.

Acknowledgements: We gratefully acknowledge Dr. Isabelle Berrebi-Bertrand and Dr. Bernard Gout for their fruitful scientific contribution.

References

[1] Scheffzek, K., Lautwein, A., Kabsch, W., Ahmadian, M.R. and Wittinghofer, A. (1996) Nature 384, 591–596.
[2] Barrett, T., Xiao, B., Dodson, E.J., Dodson, G., Ludbrook, S.B., Nurmahomed, K., Gamblin, S.J., Musacchio, A., Smerdon, S.J. and Eccleston, J.F. (1997) Nature 385, 458–461.
[3] Scheffzek, K., Ahmadian, M.R. and Wittinghofer, A. (1998) Trends Biochem. Sci. 23, 257–262.
[4] Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmüller, L., Lautwein, A., Schmitz, F. and Wittinghofer, A. (1997) Science 277, 333–338.
[5] Rittinger, K., Walker, P.A., Eccleston, J.F., Smerdon, S.J. and Gamblin, S.J. (1997) Nature 389, 758–762.
[6] Bax, B. (1998) Nature 392, 447–448.
[7] Calmels, T.P.G., Callebaut, I., Léger, I., Durand, P., Bril, A., Mornon, J.P. and Souchet, M. (1998) FEBS Lett. 426, 205–211.
[8] Rittinger, K., Taylor, W.R., Smerdon, S.J. and Gamblin, S.J. (1998) Nature 392, 448–449.
[9] Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J., Henrissat, B. and Mornon, J.P. (1997) Cell. Mol. Life Sci. 53, 621–645.
[10] Girault, J.A., Labesse, G., Mornon, J.-P. and Callebaut, I. (1999) Trends Biochem. Sci. 24, 54–57.
[11] Callebaut, I., Courvalin, J.C. and Mornon, J.P. (1999) FEBS Lett. 446, 189–193.
[12] Wojcik, J., Girault, J.A., Labesse, G., Chomilier, J., Mornon, J.P. and Callebaut, I. (1999) Biochem. Biophys. Res. Commun. 259, 113–120.
[13] Gouet, P., Courcelle, E., Stuart, D.I. and Metoz, F. (1999) Bioinformatics 15, 305–308.
[14] Poupon, A. and Mornon, J.P. (1998) Proteins 33, 329–342.
[15] Lee, B.K. and Richards, F.M. (1971) J. Mol. Biol. 112, 379–400.
[16] Kamtekar, S. and Hecht, M.H. (1995) FASEB J. 9, 1013–1022.
[17] Poupon, A. and Mornon, J.P. (1999) FEBS Lett. 452, 283–289.
[18] Bowie, J.U. (1997) Nat. Struct. Biol. 4, 915–917.
[19] Glykos, N.M., Cesareni, G. and Kokkinidis, M. (1999) Struct. Fold. Des. 7, 597–603.
[20] Blondel, A., Renaud, J.P., Fischer, S., Moras, D. and Karplus, M. (1999) J. Mol. Biol. 291, 101–115.
[21] Locher, K.P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J.P. and Moras, D. (1998) Cell 95, 771–778.
[22] Bullough, P.A., Hughson, F.M., Skehel, J.J. and Wiley, D.C. (1994) Nature 371, 37–43.