# Correspondence

## Common amino acid sequence motifs in p53, 14-3-3 and Akt protein families

Ximena Montano*

First published online 9 October 2001

The tumour suppressor p53 is a multifunctional phosphoprotein able to induce apoptosis, differentiation or growth arrest, by acting as a transcription factor (capable of recognising DNA specific sequences) and targeting expression of genes involved in cell cycle control [1,2]. p53 is a member of a family of related proteins (p63/CUSP/KET/p51/p40 and p73) characterised by multiple isoforms which share similar biological functions [3–7]. Recently, we have shown that p53 binds to trk A (the receptor for nerve growth factor (NGF)) in the presence of the non-receptor tyrosine kinase c-abl [8,9]. Importantly, p53 induces trk A hyperphosphorylation and stimulation of signalling cascades in the absence of NGF stimulation [8–10]. Therefore, because of its role in trk A signalling, the aim has been to determine whether p53 has sequence homology to proteins involved in signal transduction pathways.

From results obtained after a scan of the human p53 sequence for the presence of known motifs and, more specifically, domains common to signalling proteins using the program GAP from the GCG package (Wisconsin Package Version 9.1, Genetics Computer Group, Madison WI, USA), it was noticed that p53 and c-Akt had sequence similarity. Then the p53 and c-Akt sequences were compared using a Gibbs sampler approach [11] as implemented in the program MACAW [12]. The search parameters were: minimum pattern length 4, maximum pattern length 20, random seed 12345, number of trials six, iterations per 50. The scoring matrix was BLOSUM62. The comparison revealed two regions of similarity, D1 and D2, with segment-pair scores of 1193 and 1353, and a $P$ value (for each domain) of less than $10^{-20}$ respectively; thus indicating that the relationship was significant. To detect other proteins containing these motifs, D1 or D2 of the human p53 sequence were subsequently used as input to a FASTA database search [13]. The FASTA search was performed using the BLOSUM50 matrix and a $k$-tuple value of 2. However, despite the MACAW results, this search showed that the presence of either motif alone was not highly significant. This result was expected as D1 and D2 are short peptide regions. Thus, the possibility of proteins containing both motifs was investigated with a low-stringency search using the FINDPATTERNS program of the GCG package. The sequence pattern syntax covered both D1 and D2 and was flexible enough to allow detection of related homologues (see Fig. 1). Results revealed that the combined presence of D1 and D2 is a feature of the p53, 14-3-3 and Akt/PKB/RAC protein families. The database used was a non-redundant protein sequence dataset compiled from all existing public protein databases to date.

Fig. 1A shows the consensus amino acid sequence homology of representative members of the p53 family of proteins with representative members of the 14-3-3 and Akt/PKB/RAC families respectively. D1 and D2 are separated by a non-homologous linker region, but in the case of 14-3-3 proteins the order of these domains is reversed (Fig. 1B). This is the first analysis showing shared amino acid sequences between these families.

The transactivation domain, DNA binding core and carboxyl-terminus of human p53 lie within amino acid residues 1–43, 100–300 and 301–393 respectively. D1 and D2 are located in the DNA binding core of the p53 family, with D1 in the C-terminus of conserved region II, a region prone to germline mutations [14], and D2 proximal to conserved region III. Importantly, the spatial arrangement of D1 and D2 in the crystal structure of DNA-bound p53 shows these regions not bound to DNA and accessible to protein–protein interactions (Fig. 1C, upper panels) [15].

Heat shock proteins/molecular chaperones (HSPs) have several functions; these include stabilisation of newly synthesised peptides, translocation of peptide chains across membranes and assisting assembly or disassembly of multimeric complexes [16,17]. Members of the HSP70 family of molecular chaperones such as Hsc70 and HSP90 have been shown to associate with mutant rather than wild type p53 [18–20]. However, more recently, when experiments were carried out using a library of wild type p53 overlapping peptides, it was shown that sequences of the DNA core domain were able to bind Hsc70 and peptides containing D1 and D2 bound with high affinity [21]. These data suggest that the biological function of D1 and D2 is to mediate p53 and Hsc70 association.

Akt/PKB/RAC proteins are cytoplasmic serine/threonine kinases and direct targets of growth factor or insulin-activated phosphatidylinositol (PI)-3 kinase [22,23]. These proteins contain an amino-terminal pleckstrin homology (PH) domain and a catalytic kinase domain (residues 1–109 and 148–411 respectively) linked by a highly acidic region (residues 110–147). The PH domain of Akt is involved in protein interactions and phospholipid binding [24]. Crystallographic data of the PH domain from pleckstrin N show this domain exposed within the molecule [25]. In Akt D1 is located in the C-terminus of the PH domain and D2 in the C-terminus of the acidic linker. Unfortunately, the crystal structure of Akt has not been solved; thus, it is not possible to define the spatial arrangement of D1 and D2 in this protein. There is evidence showing that growth factor regulation of Akt can be PI-3 kinase-dependent or -independent [22,26]. Proteins such as v-src and sevenless are regulated by associating to a complex of cdc37 and HSP90 [27]. These complexes seem to act as general chaperones for protein kinases [27]. Interestingly, there is evidence suggesting that Akt is also regulated by this process and that it can directly bind to HSP90 [26,28]. Importantly, the presence of D1 and D2 suggests that Hsc70 could potentially associate and be involved in Akt regulation.

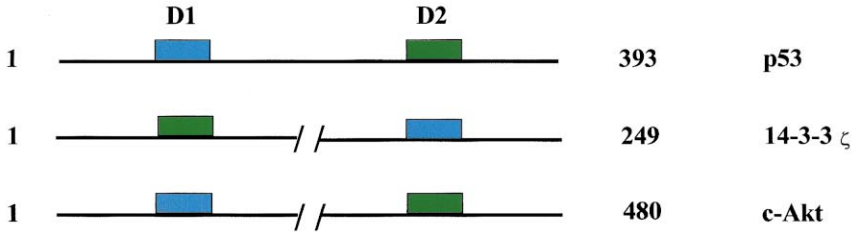
14-3-3 proteins are a family of highly conserved molecules. They can regulate and interact with phosphoserine containing proteins known to participate in signalling pathways involved in the control of cell proliferation, differentiation and trans-


PII: S0014-5793(01)02903-9

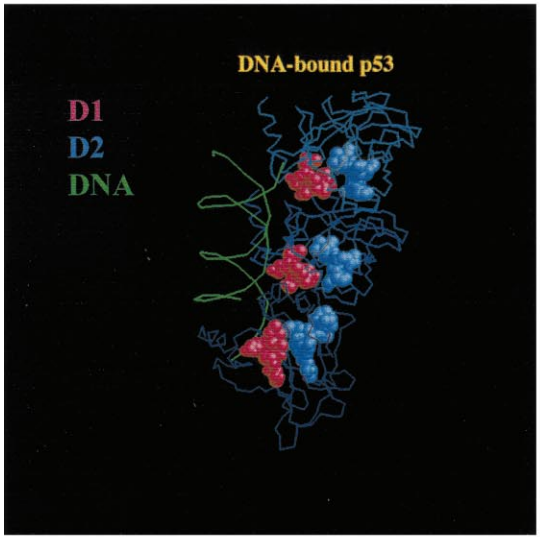
A

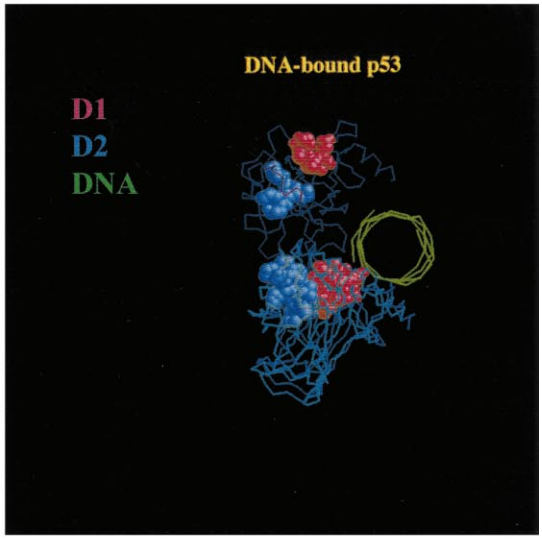
| | | D1 (1 6) | D2 (1 8) | | |
|---|---|---|---|---|---|
| p53 Human | N | CQLAKT | PPQHLIRV | C | 135-140/190-197 |
| p53 Mouse | N | CQLAKT | PPQHLIRV | C | 132-137/187-194 |
| p73 α, β Human | N | CQIAKT | PASHLIRV | C | 153-158/210-217 |
| p63 α, β, γ Human | N | CQIAKT | PPSHLIRV | C | 164-169/221-228 |
| 14-3-3ζ Human | C | CSLAKT | P THPIRL | N | 189-194/162-168 |
| 14-3-3ζ Drosophila | C | CQLAKQ | P THPIRL | N | 192-197/165-171 |
| Rad24 S. pombe | C | CYLAKQ | P THPIRL | N | 194-199/167-173 |
| v-Akt | N | CQLMKT | P KH RV | C | 82-87/164-168 |
| c-Akt Human | N | CQLMKT | P KH RV | C | 60-65/141-145 |
| c-Akt Chicken | N | CQLMKT | P KH KV | C | 60-65/141-145 |

B




C



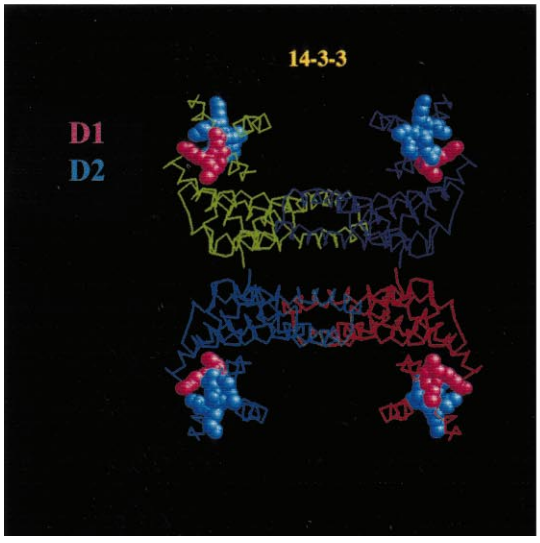




Fig. 1. Multiple sequence alignment of representative members of p53, 14-3-3 and Akt families. A: Alignment showing amino acid residues of D1 and D2. Because of the large number of proteins belonging to each family, especially 14-3-3, only a set of representative members are depicted. D1 and D2 were aligned using the MACAW multiple sequence alignment program [12]. All conserved residues are shown in orange. In D1 A at position 4 is conserved between p53 and 14-3-3 families and T at position 6 is conserved in p53 and Akt members (shown in yellow). All proteins with exception of 14-3-3ζ and Rad24 have Q at position 2 (shown in yellow) which is replaced by S and Y respectively (shown in white). Interestingly, comparative analyses of all 14-3-3 family members to date show high amino acid variability at D1 position 2 containing residues N, H, S, Q, L, T, G, H, D or Y (from which N is the most prevalent and Y is the least common) and low variability at amino acid position 6 containing either Q or T (of which Q is the most prevalent) (data not shown). In D2 I at position 6 is conserved between members of p53 and 14-3-3 and V at position 8 is conserved between members of p53 and Akt (shown in yellow) whereas R at position 7 is conserved between members of p53, 14-3-3 and Akt (shown in orange) but is replaced by K in chicken c-Akt (shown in yellow). Comparative analyses of all v- and c-Akt members to date showed only chicken c-Akt containing K at position 8. L and P at position 5 are conserved within the p53 and 14-3-3 families respectively (shown in yellow and white). P at position 2 is conserved in all p53 proteins and in members of p63 (shown in white), but not in p73α and β in which it is replaced by A. T and K at position 3 are conserved within 14-3-3 and Akt families respectively (shown in white). Q at position 3 is conserved within p53 proteins and S is conserved within members of p63 and p73. Importantly, comparative analysis of all 14-3-3 proteins to date shows D2 containing the same amino acid sequences in all members of the family. The pattern for FINDPATTERNS of the GCG program used for the searches is:

$$Cx(L,I)(A,M)K(T,Q)\times\{15,75\}Px\{0,2\}Hx\{0,2\}(R,K)(V,I,L)$$

$$Px\{0,2\}Hx\{0,2\}(R,K)(V,I,L)\times\{15,75\}Cx(L,I)(A,M)K(T,Q)$$

The Ensembl project (http://www.ensembl.org/; [35]) produces automated annotation of the publicly available human genomic sequence data by using gene prediction software and a variety of analytical tools. The predictions are split into confirmed and predicted genes (and their translated products) on the basis of supporting evidence from similarity to known ESTs, genes and proteins in the public sequence databases. The predicted and confirmed protein datasets were also used (as already described) to detect D1 and D2. Interestingly, results showed that in both the predicted as well as in the confirmed sets D1 and D2 were present in 14-3-3 as well as p53 and p73α and β members. The presence of c-Akt was not seen in current searches; this is probably due to the lack of data, to date, available from the Ensembl project; this most likely will be obtainable in the future while progress of the human genome analysis is made. Accession numbers: human p53 P04637; mouse p53 P02340; human p73α,β 015350; human p63α, β and γ AAC62635, AAC62637, AAC62633; human 14-3-3ζ P29312; *Drosophila* 14-3-3ζ P29310; *Schizosaccharomyces pombe* Rad24 P42656; v-Akt P31748; human c-Akt P31749; chicken c-Akt AAB94767. B: Schematic representation of D1 (blue) and D2 (green). C: Stereoview of the spatial arrangement of D1 (red) and D2 (blue) showing two different orientations in DNA (green)-bound p53 (upper panels) and in 14-3-3ζ- or raf (yellow)-bound 14-3-3ζ (lower panels) using Rasmol software. PDB identifier numbers: DNA-bound p53: 1TSR; 14-3-3ζ dimer: 1A40; raf-bound 14-3-3ζ: 1A37.
←

formation [29,30]. The monomer contains nine α-helices and regions with homology to the PKC pseudosubstrate and annexin respectively (amino acids 53–56 and 122–136 in 14-3-3ζ). The crystal structure was solved as a dimer revealing an amphipathic groove that mediates c-raf binding [30,31]. Spatial analysis of 14-3-3ζ shows D1 and D2 adjacent to this groove (Fig. 1C, lower left panel). Although amino acids from this groove are directly involved in several protein interactions, there is evidence that residues spanning amino acids 171–213 (termed box 1) can participate in protein–protein associations [30,32–34]. For example, a sequence containing D2 (amino acids 163–187) contributes to c-raf recognition [33] (Fig. 1C, lower right panel). Proteins such as platelet glycoprotein 1bx bind to a region containing D1 (amino acids 188–209) [33] and tryptophan hydroxylase can associate with the box 1 region which contains both D1 and D2 [34]. Importantly, computer analysis shows that D1 and D2 are greatly conserved among all members of the family, suggesting that both motifs, in all 14-3-3 proteins, appear to be involved in protein–protein interactions. To date there is no evidence showing 14-3-3 binding to HSP70 members, but the presence of these motifs suggests that it is possible that HSP70 molecules could compete for binding with, for example, the above proteins at the D1 and D2 sites.

In conclusion these data show that there are two regions within members of p53, 14-3-3 and Akt protein families which may be evolutionarily conserved. Their presence possibly leads to a common biological function such as regulation and binding to heat shock chaperones.

*Acknowledgements:* Thanks to J. Sgouros, G. Smythers and especially to M. Mitchell for assistance with sequence analysis programs, R. Liddington for access to the 14-3-3 crystallographic coordinates, C. Dickson for critical reading of this manuscript and D. Morrison for useful discussions.

## References

[1] Prives, C. and Hall, P.A. (1999) J. Pathol. 187, 112–126.
[2] Giaccia, A.J. and Kastan, M.B. (1998) Genes Dev. 12, 2973–2983.
[3] Kaghad, M., Bonnet, H., Yang, A., Creancier, L., Biscan, L.C., Valent, A., Minty, A., Chalon, P., Lelias, J.-M., Dumont, X., Ferrara, P., McKeon, F. and Caput, D. (1997) Cell 90, 809–819.
[4] Schmale, H. and Bamberger, C. (1997) Oncogene 15, 1363–1367.
[5] Osada, M., Ohba, M., Kanahara, C., Ishioka, C., Kanamuru, R., Katoh, I., Ikawa, Y., Nimura, Y., Nagakawara, A., Obinata, M. and Ikawa, S. (1998) Nat. Med. 4, 839–843.
[6] Yang, A., Kaghad, M., Wang, Y., Gillet, E., Fleming, M.D., Dotsch, V., Andrews, C.N., Caput, D. and McKeon, F. (1998) Mol. Cell 2, 305–316.
[7] Strano, S., Rossi, M., Fontemaggi, G., Munarriz, E., Soddu, S., Sacchi, A. and Blandino, G. (2001) FEBS Lett. 490, 163–170.
[8] Montano, X. (1997) Oncogene 15, 245–256.
[9] Brown, A., Browes, C., Mitchell, M. and Montano, X. (2000) Oncogene 19, 3032–3040.
[10] Browes, C., Rowe, J., Brown, A. and Montano, X. (2001) FEBS Lett. 497, 20–25.
[11] Lawrence, C.E., Altschul, S.F., Boguski, M.S., Liu, J.S., Neuwald, A.F. and Wootton, J.C. (1993) Science 262, 208–214.
[12] Schuler, G.D., Altschul, S.F. and Lipman, D.J. (1991) Proteins 9, 180–190.
[13] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[14] Hollstein, M., Rice, K., Greenblatt, M.S., Soussi, T., Fuchs, R., Soilie, T., Howig, E., Smith Sorensen, B., Montesano, R. and Harris, C.C. et al. (1994) Nucleic Acids Res. 22, 3551–3555.
[15] Arrowsmith, C.H. and Morin, P. (1996) Oncogene 12, 1379–1385.
[16] Gething, M.J. and Sambrook, J.F. (1992) Nature 355, 33–45.
[17] Helmbrecht, K., Zeise, E. and Rensing, L. (2000) Cell Prolif. 33, 341–365.

[18] Hinds, P.W., Finlay, C.A., Frey, A.B. and Levine, A.J. (1987) Mol. Cell Biol. 7, 2863–2869.
[19] Sepehrnia, B., Benjamin Paz, I., Dasgupta, G. and Momand, J. (1996) J. Biol. Chem. 271, 15084–15090.
[20] Whitesell, L., Sutphin, P.D., Pulcini, E.J., Martinez, J. and Cook, P.H. (1998) Mol. Cell Biol. 18, 1517–1524.
[21] Fourie, A.M., Hupp, T.R., Lane, D.P., Sang, B.-C., Barbosa, M.S., Sambrook, J.F. and Gething, M.-J.H. (1997) J. Biol. Chem. 272, 19471–19479.
[22] Datta, S.R., Brunet, A. and Greenberg, M.E. (1999) Genes Dev. 13, 2905–2927.
[23] Blume-Jensen, P. and Hunter, T. (2001) Nature 411, 355–365.
[24] Coffer, P.J., Jin, J. and Woodgett, J.R. (1998) Biochem. J. 335, 1–13.
[25] Yoon, H.S., Hadjduk, P.J., Petros, A.M., Olejniczak, E.T., Meadows, R.P. and Fesik, S.W. (1994) Nature 369, 672–675.
[26] Bellacosa, A., Chan, T., Ahmed, N.N., Datta, K., Malstrom, S., Stokoe, D., McCormick, F., Feng, J. and Tsichlis, P. (1998) Oncogene 17, 313–325.
[27] Hunter, T. and Poon, Y.C. (1997) Trends Cell Biol. 7, 157–161.
[28] Sato, S., Fujita, N. and Tsuro, T. (2000) Proc. Natl. Acad. Sci. USA 97, 10832–10837.
[29] Reuther, G.W. and Pendergast, A.M. (1996) Vitam. Horm. Adv. Res. Appl. 52, 149–175.
[30] Fu, H., Subramanian, R.R. and Masters, S.C. (2000) Annu. Rev. Pharmacol. Toxicol. 40, 617–647.
[31] Dubois, T., Howell, S., Amess, B., Kerai, P., Learmonth, M., Madrazo, J., Chandhri, M., Rittinger, K., Scarabel, M., Soneji, Y. and Aitken, A. (1997) J. Prot. Chem. 16, 513–522.
[32] Petosa, C., Masters, S.C., Pohl, J., Bankston, L.A., Wang, B., Fu, H. and Liddington, R. (1998) J. Biol. Chem. 273, 16305–16310.
[33] Gu, M. and Du, X. (1998) J. Biol. Chem. 273, 33465–33471.
[34] Ichimura, T., Uchiyawa, J., Kunikiro, O., Ito, M., Horigone, T., Omata, S., Shinkai, F., Koji, H. and Isobe, T. (1995) J. Biol. Chem. 270, 28515–28518.
[35] International Human Genome Sequencing Consortium (2001) Nature 409, 860–921.

*Present address: Department of Biology, Imperial College of Science Technology and Medicine, Sir Alexander Fleming Building, Exhibition Road, London SW7 2AZ, UK. Fax: (44)-020-7584 2056.
E-mail address: x.montano@ic.ac.uk (X. Montano).

*Cancer Research Unit, Medical School, Framlington Place, University of Newcastle, Newcastle NE2 4HH, UK*

PII: S0014-5793(01)02903-9