The EMBO Journal vol.11 no.9 pp.3263-3269, 1992

The human poly(ADP-ribose) polymerase nuclear localization signal is a bipartite element functionally separate from DNA binding and catalytic activity

Valérie Schreiber, Miguel Molinete, Hélène Boeuf[1], Gilbert de Murcia and Josiane Ménissier-de Murcia[2]

Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique associé à L'Université Louis Pasteur, UPR de Cancérogénèse et de Mutagénèse Moléculaire et Structurale, 15 rue René Descartes, F-67084 Strasbourg Cedex and [1]Laboratoire de Génétique Moléculaire des Eucaryotes, 11 rue Human, F-67085 Strasbourg Cedex, France
[2]Corresponding author

Communicated by R.A.Laskey

Poly(ADP-ribose) polymerase (PARP, EC 2.4.2.30) is a zinc finger DNA-binding protein involved in DNA repair processes in eukaryotes. By deletion and extensive site-directed mutagenesis, its DNA-binding domain fused to the N-terminus of β-galactosidase was shown to contain a nuclear localization signal (NLS) of the form KRK-X(11)-KKKSKK (residues 207-226). *In vitro*, both the DNA-binding capacity and the polymerizing activity of PARP are independent of the nuclear location function. Each basic cluster is essential but not sufficient on its own for this function, while both motifs together are. Crucial basic amino acids (K207, R208 and K222) in each of these two motifs are required for nuclear homing. The results presented here support the concept that the human PARP NLS is an autonomous functional element and belongs to the class of bipartite NLSs. We show that the linear distance between the two basic clusters is not crucial. Insertional mutation analysis leading to a partial reversion of the cytoplasmic phenotype displayed by the mutant K222I highlights the crucial positioning of this lysine. The structure-function relationship of the second cluster of basic residues is discussed.

Key words: β-galactosidase fusion protein/eukaryotic expression vector/mutagenesis/nuclear transport/zinc finger protein

Introduction

In eukaryotic cells, the nuclear membrane separates the genome and its associated functions from the cytoplasm. Following synthesis, proteins involved in the management of the genes must enter the nucleus through nuclear pore complexes which constitute the morphological superstructure mediating exchanges between nucleus and cytoplasm (Feldherr *et al.*, 1984; Stewart *et al.*, 1990). Proteins containing a nuclear localization signal (NLS) are actively translocated into the nucleus (Dingwall and Laskey, 1986; Newmeyer *et al.*, 1986; Dingwall, 1991). The nuclear residency of proteins could be the result of a shuttle mechanism, as shown recently for the progesterone receptor (Guiochon-Mantel *et al.*, 1991), the nuclear proteins diffusing into the cytoplasm and being constantly transported back into the nucleus. The first evidence for the existence of NLS was initially provided by Dingwall *et al.* (1982), who demonstrated by deletion analysis that such a karyophilic signal is present in the C-terminal part of the nucleoplasmin, a major nuclear protein of *Xenopus* oocytes.

Many NLSs have been identified so far (Roberts, 1989; Dingwall and Laskey, 1991; Garcia Bustos *et al.*, 1991; Schreiber *et al.*, 1992); the prototype, that of SV40 large T-antigen, is the sequence $_{126}$PKKKRKV$_{132}$, which is necessary and sufficient to promote nuclear accumulation of a reporter protein (Kalderon *et al.*, 1984). A single amino acid change (K128 to T or N) causes the normally nuclear protein to be cytoplasmically localized (Kalderon *et al.*, 1984; Landford and Buttel, 1984; Landford *et al.*, 1986). Furthermore, the region 111-125, which flanks the nuclear homing signal and contains an *in vivo* casein kinase II site, dramatically enhances the rate of nuclear uptake (Rihs and Peters, 1989; Rihs *et al.*, 1991).

No consensus sequence emerges among the identified NLSs, suggesting that shape and charge rather than amino acid sequence might be of primary importance for signal recognition. However, it appears that they are typically (i) short basic amino acids stretches, (ii) exposed on the surface of the protein in order to be easily recognized by a receptor protein, (iii) not removed following translocation as are signal peptides and (iv) single or multiple in a given protein.

So far, NLSs can be roughly partitioned into two classes: (i) the SV40 large T-like NLSs which are one compact sequence and (ii) the bipartite signals, recently described by Robbins *et al.* (1991) which encompass two interdependent basic domains. Interestingly, this latter signal could be widespread with regard to alignment of candidate sequences of nuclear proteins.

Two approaches have been used to identify NLS sequences in proteins. In the first one, the subtractive approach, the karyophilic signal is assessed in deletion or point mutants defective in nuclear entry. The second strategy, the additive approach, seeks to determine which part of a nuclear protein is sufficient to promote nuclear accumulation of a reporter protein.

We have combined these two approaches to define precisely the NLS of the human poly(ADP-ribose) polymerase (PARP, EC 2.4.2.30). This chromatin associated protein catalyses the covalent attachment of ADP-ribose units from the substrate NAD^+ to various nuclear acceptor proteins (for a review see Althaus and Richter, 1987; de Murcia *et al.*, 1991). It has been postulated that this post-translational modification influences several chromatin functions, especially those involving nicking and rejoining of DNA strands, such as cell proliferation, differentiation and recovery from DNA damage. PARP (113.6 kDa) is a multifunctional enzyme (see Figure 1A) and its activity is strongly stimulated by DNA strand breaks (Benjamin and Gill, 1980). Following limited proteolysis of the purified enzyme, three functional domains have been identified (Kameshita *et al.*, 1984): an N-terminal 46 kDa fragment

including a zinc finger DNA-binding domain which directs the specific recognition of single strand breaks (Mazen *et al.*, 1989; Ménissier de Murcia *et al.*, 1989; Gradwohl *et al.*, 1990), a central 22 kDa polypeptide containing the automodification sites and a C-terminal fragment (54 kDa) bearing the catalytic domain (Simonin *et al.*, 1990).

In this study, we show that (i) the sequence required to promote nuclear accumulation of the human PARP is located in the DNA-binding domain and this targeting function is totally independent of both the DNA binding capacity and the polymerizing activity of PARP, (ii) this nuclear location signal encompasses two basic motifs related to the bipartite NLSs (Robbins *et al.*, 1991), (iii) each of these two motifs is necessary but not sufficient to target β-galactosidase into the nucleus, but both motifs together are and (iv) an increased or decreased distance between the two clusters is tolerated but the correct positioning of a basic residue in the second motif is required.

Results

Sequence analysis

Although there is no consensus sequence among the many NLSs that have been described so far, there are some features indicating that a given sequence could act as a targeting signal; in most cases the motif $K^K/_R X^K/_R$ is present in karyophilic signals (Chelsky *et al.*, 1989).

Sequence analysis of the entire PARP revealed three motifs rich in basic amino acid matching the criteria for a nuclear targeting signal (Figure 1B). The first motif (amino acids 207–226) is localized in the 29 kDa N-terminal DNA-binding domain. It is composed of two basic elements, $_{207}$KRK-X(11)-KKKSKK$_{226}$, which resemble the bipartite NLS described by Robbins *et al.* (1991). Furthermore, it is well conserved throughout evolution, although the mouse and chicken sequences displayed a conservative change: an arginine instead of a lysine at position 225. In the second motif, $_{346}$KKLKVKK$_{352}$, only the basic residues are conserved. On the other hand, the third motif, $_{505}$KKSK$_{508}$, completely diverges both in the chicken and in the *Xenopus* sequences, indicating that this motif might not be important.

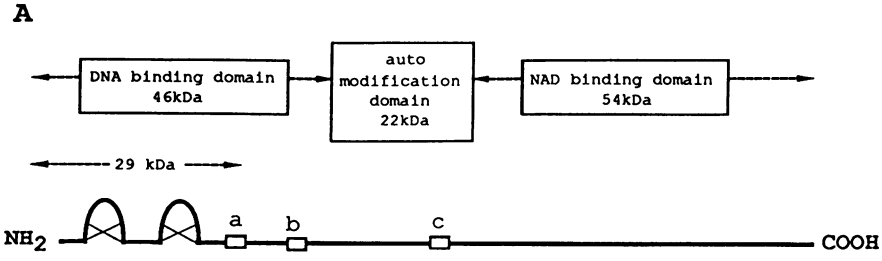
Experimental strategy

The sequences encoding the wild-type or the mutated N-terminal part of PARP were cloned in-frame with the β-galactosidase-encoding *lacZ* gene of the pCH110 vector (Hall *et al.*, 1983). Using a histochemical stain for β-galactosidase activity (Sanes *et al.*, 1986), the approach was to assess the subcellular distribution of the chimeric proteins transiently expressed from recombinant plasmid in transfected HeLa cells. In a transfected culture only some of the cells appeared to take up and express the exogenous DNA as judged by their blue staining; other cells remained relatively free of staining. This provided a useful internal control for the specificity of the reaction with X-Gal.

When HeLa cells were transfected with the vector pCH110, the cells were stained blue, indicating that β-galactosidase activity was located in the cytoplasm (data not shown). As a positive control, the SV40 large T-antigen NLS sequence ($_{125}$PPKKKRKVEDPV$_{136}$, Kalderon *et al.*, 1984) was cloned into the *Kpn*I site of the pCH110 vector in-frame with the *lacZ* gene, resulting in pCHSV. Cells transfected with this plasmid stained blue only in the nucleus, demonstrating that β-galactosidase was targeted into nuclei (data not shown).

A nuclear location signal is present within the 29 kDa DNA-binding domain

We first focused on the putative bipartite NLS localized in the 29 kDa N-terminal DNA-binding domain. The cDNA encoding the N-terminal fragment (residues 1–233) was cloned into the *Kpn*I site of pCH110 using an adaptor in-frame with the *lacZ* gene, resulting in pCHNPARP. The location of the β-galactosidase fusion protein expressed in HeLa cells transfected with pCHNPARP was exclusively nuclear, consistent with the view that a nuclear targeting



B

		a		b		c
hPARP	:	(207) KRK-X(11)-KKKSKK	(346)	KKLKVKK	(505)	KKSK
mPARP	:	(207) KRK-X(11)-KKKSRK	(346)	KKLKVKK	(504)	KKSK
bPARP	:	(210) KRK-X(11)-KKKSKK	(349)	KKLKIKK	(507)	KKSK
cPARP	:	(207) KRK-X(10)-KKKSRK	(345)	KKFKCKK	(501)	MQSA
xPARP	:	(207) KRK-X(11)-KKKIKK	(345)	KKFKFKR	(505)	GKSS

Fig. 1. (**A**) Schematic representation of the different functional domains of poly(ADP-ribose) polymerase. The two zinc fingers are represented as loops joined by an X. The three basic motifs a, b and c, which resemble karyophilic signals, are boxed. (**B**) Positions of the basic motifs from the human (h), murine (m), bovine (b), chicken (c) and *Xenopus* (x) poly(ADP-ribose) polymerases. The sequences were taken respectively from Uchida *et al.* (1987), Huppi *et al.* (1989), Saito *et al.* (1990), Ittel *et al.* (1991) and B.Saulier (personal communication).

sequence resides in the N-terminal part of the protein (Figure 4).

To identify the function of this sequence, our strategy was to delete the basic core, $_{221}$KKKSKK$_{226}$, resulting in the plasmid pCHNPARPΔ2 and to study the resulting alteration in the subcellular localization of the encoded chimeric protein. In HeLa cells transfected with pCHNPARPΔ2, the chimeric protein remained cytoplasmic, indicating that the deleted sequence is required to transport the protein into the nucleus (Figure 4).

Since the efficiency of an NLS is sensitive to the protein context within which it is present (e.g. SV40 large T-antigen NLS becomes inactive when placed in a hydrophobic region of pyruvate kinase (Roberts *et al.*, 1987), the significance of this sequence was investigated in its natural protein environment. The subcellular distribution of either the wild-type PARP or a mutant with the six amino acids $_{221}$KKKSKK$_{226}$ deleted was determined by immunofluorescence staining of the expressed protein using a polyclonal antibody raised against the second zinc finger (Simonin *et al.*, 1991). We first confirmed the nuclear localization of the wild-type PARP protein expressed in HeLa cells transfected with pECVPARP (Figure 2a). In contrast, deletion of the basic core $_{221}$KKKSKK$_{226}$ yielded a mutant protein which accumulated in the cytoplasm of cells transfected with pECVPARPΔ2 (Figure 2b). These results indicate that the deleted sequence is necessary for the nuclear targeting function and that the two other basic stretches $_{346}$KKLKVKK$_{354}$ and $_{505}$KKSK$_{508}$ cannot promote the nuclear accumulation of the mutant PARPΔ2. These results must be related to the fact that PARP contains only one NLS located within the 29 kDa DNA-binding domain and that the hexapeptide $_{221}$KKKSKK$_{226}$ is at least part of it.

The NLS function is involved neither in DNA-binding nor in poly(ADP-ribose) polymerase activity

Examination of the PARP nuclear targeting sequence suggests that it might bind to DNA due to its positive charges and, since the DNA-binding capacity governs the enzymatic activity of PARP, removal of the NLS sequence might also affect this function. To test this hypothesis, both the wild-type enzyme and the NLS-deleted version were overexpressed in bacteria, blotted onto nitrocellulose sheets and probed with radiolabelled nicked DNA or incubated with [^{32}P]NAD$^+$. *Escherichia coli* TGE900 cells transformed with the plasmids pTGPARP and pTGPARPΔ2 express, after induction by a temperature shift, polypeptides of 116 kDa (Figure 3, lanes b and c) which can be identified as PARP polypeptides with our set of poly- and monoclonal antibodies (data not shown). The DNA binding capacity of the two recombinant proteins was determined by South-western blotting (Mazen *et al.*, 1989; Ménissier de Murcia *et al.*, 1989) (Figure 3, lanes d and e). The poly(ADP-ribose) polymerase activity was determined by the 'activity blot' technique (Simonin *et al.*, 1991), which allows the detection of poly(ADP-ribose) synthesis from [^{32}P]NAD$^+$, of polypeptides transferred to a nitrocellulose sheet (Figure 3, lanes f and g). Using these two methods, we could no longer monitor any difference between wild-type and deleted proteins, thus demonstrating that the basic sequence $_{221}$KKKSKK$_{226}$ does not contribute to the DNA-binding capacity of PARP.

Crucial amino acids involved in targeting function

To investigate the detailed amino acid requirement for nuclear accumulation, deletions and point mutations were introduced by site-directed mutagenesis into the 29 kDa DNA-binding domain (Figure 4). All mutated peptides were transiently expressed in-frame with β-galactosidase in HeLa cells.

Both deleted proteins (NPARPΔ1 and NPARPΔ2, see above) were entirely excluded from the nucleus (Figure 4B) showing the fundamental importance of each ot the two basic elements constituting the NLS.

Point mutations converting basic residues to amino acids with equivalent volumes had different effects. Whereas K209I, K221I, K223I, K225I, K226I and K230I constitute a class of mutants whose phenotype was similar to that of the wild-type protein (Figure 4A), both mutants R208Q and K222I totally lost their ability to accumulate in nuclei (Figure 4B). Furthermore, the change of lysine 207 to isoleucine (mutant K207I) modulated but did not abolish

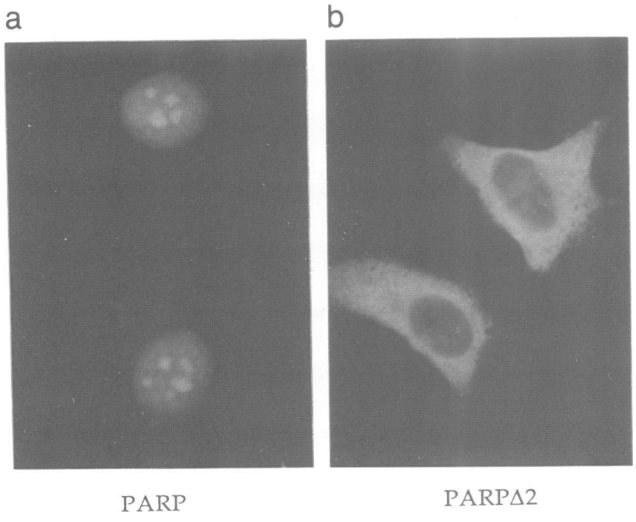


Fig. 2. Immunofluorescence detection of overexpressed PARP and PARPΔ2 proteins in transfected HeLa cells. The subcellular distribution of the overexpressed proteins was determined *in situ* by staining with a polyclonal antibody raised against the second zinc finger of PARP (Simonin *et al.*, 1991) and rhodamine (TRITC) conjugated goat anti-rabbit IgG. The vectors used were (**a**) pECVPARP and (**b**) pECVPARPΔ2.

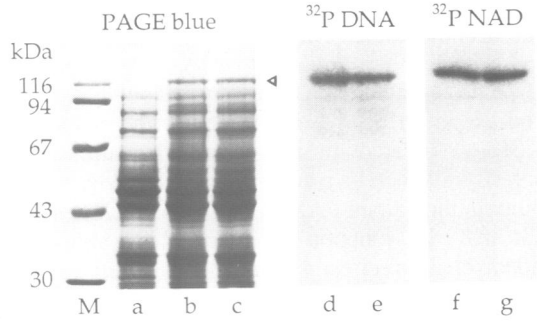


Fig. 3. DNA binding capacity and activity of the wild-type and the deletion-bearing mutant PARPΔ2 by South-western blotting and by 'activity blotting'. Crude extracts of vector containing TGE 900 bacteria were analysed by SDS–PAGE. a: pTGPARP non-induced; b, d and f: pTGPARP induced; c, e and g: pTGPARPΔ2 induced; a, b and c: Coomassie-stained gel; d and e: South-western blot; f and g: activity blot. M: molecular mass marker. The arrowhead marks the position of the expressed proteins.

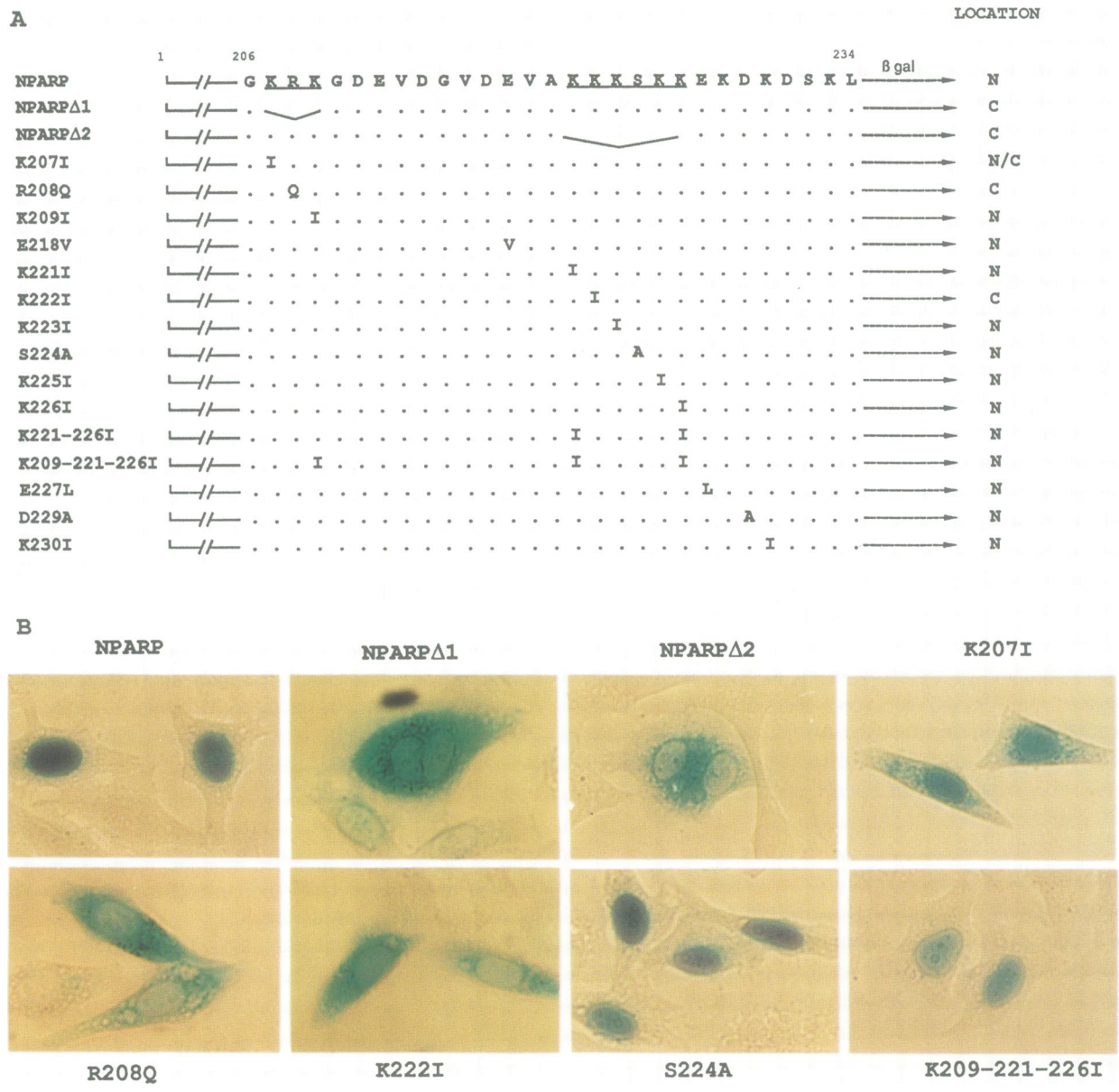


Fig. 4. Mutagenesis of the PARP nuclear localization signal (subtractive approach). (**A**) The PARP NLS is given in the single-letter code (residues 206–231) at the top. Mutants are named in the left panel and their subcellular localization is shown in the right panel (N, nuclear; C, cytoplasmic; N/C, nuclear and cytoplasmic). (**B**) Histochemical staining micrographs are shown for selected mutations as indicated.

nuclear accumulation. The double mutant K221I K226I showed a nuclear accumulation comparable to that of the wild-type protein while the remaining second basic motif is reduced to $_{222}$KKSK$_{225}$. In contrast, the same motif, $_{505}$KKSK$_{508}$, located in the automodification domain (Figure 1B) failed on its own to restore the nuclear targeting capability in the mutant PARPΔ2 (see above). These results suggest that the function of lysine 222, although not dependent on a precise environment of positive charges, requires an appropriate spatial positioning to be recognized by the nuclear transport machinery. Furthermore, the triple mutant K209I K221I K226I accumulated in the nucleus, indicating that the sequence $_{207}$KR-X(13)-KKSK$_{225}$ is sufficient for the nuclear targeting function of PARP.

It is interesting to note that numerous acidic residues surround the basic domains, which seems to be a general feature of these signals. This suggests that glutamate and aspartate could be involved in neutralizing the positive charges of lysine and arginine and consequently contribute to the stabilization of the three-dimensional structure. In order to clarify this question, aspartate 218 and 227 and glutamate 229 residues were substituted for valine, leucine and alanine respectively (mutants E218V, E227L and D229A). None of these single mutations altered the nuclear targeting capability of recombinant proteins (Figure 4), indicating an apparent lack of importance of these acidic residues.

The two basic motifs are both required for the nuclear targeting function

To highlight the central importance of the cooperation between the two basic motifs in promoting the nuclear accumulation of PARP, we generated two constructs expressing separately the two basic elements in-frame with β-galactosidase giving rise to pCHD1 and pCHD2 respectively. As expected, neither of these sequences was

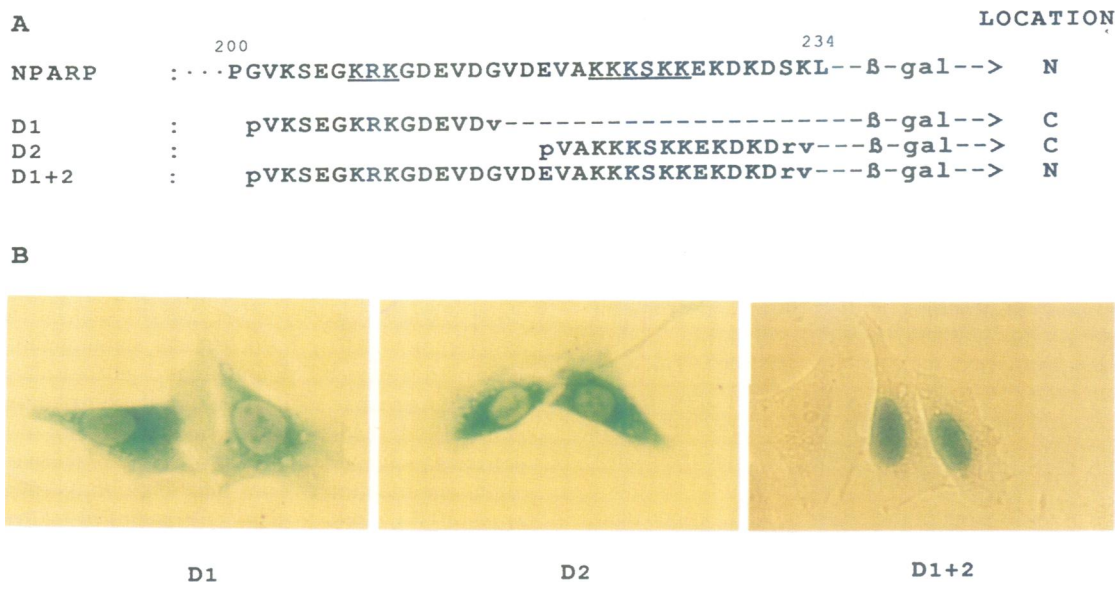


Fig. 5. Sequence requirement of the PARP nuclear location signal (additive approach). (**A**) The two basic domains (residues 202–214 and 219–231 respectively) in fusion with β-galactosidase were separately expressed in Hela cells from the constructs pCHD1 and pCHD2. Both domains in fusion with β-galactosidase were expressed in cells transfected by pCHD1+2. The right panel shows the subcellular localization of the chimeric proteins. (**B**) Histochemical staining micrographs are shown for selected mutations as indicated.

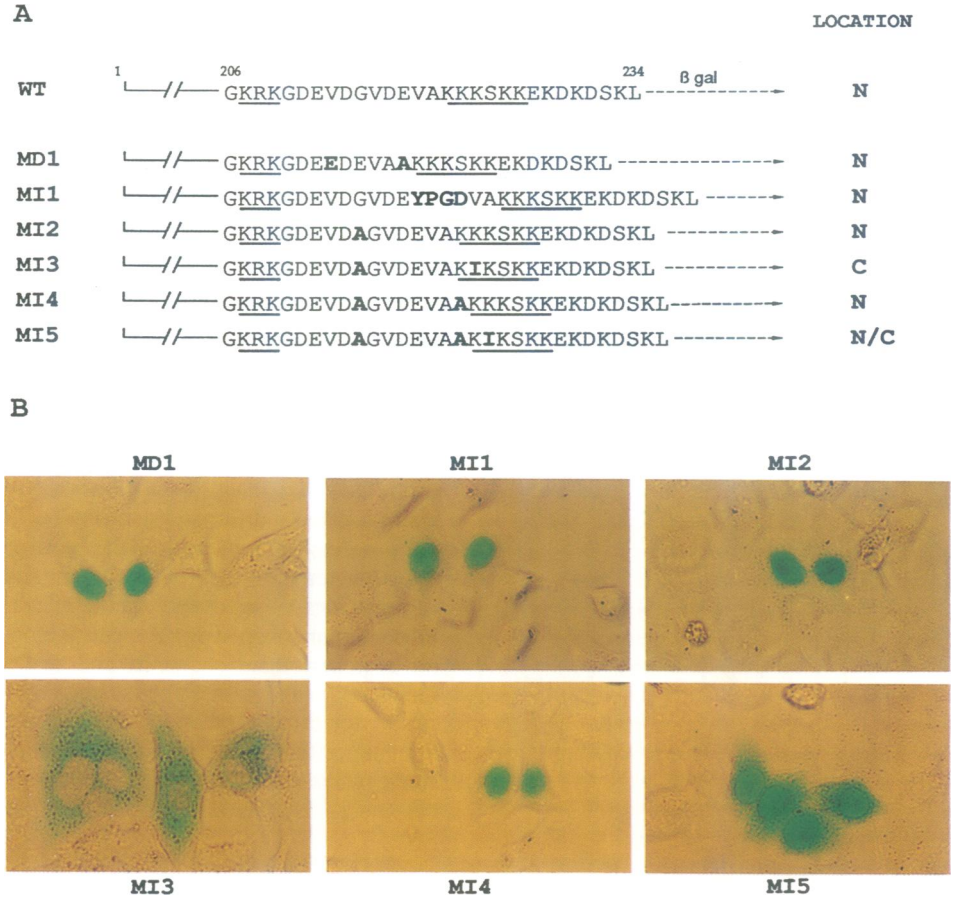


Fig. 6. Deletion and insertion mutants in the PARP nuclear location signal. (**A**) MD1 is a mutant with a two amino acid deletion; MI1 is an insertional mutant in which a β-turn has been introduced (Dyson *et al.*, 1988). MI2 and MI4 are insertional mutants conserving the wild-type second basic cluster. MI3 and MI5 possess one or two residues inserted into the K222I mutant. The inserted amino acids are shown in bold. The two basic clusters are underlined. (**B**) Histochemical staining micrographs are shown for the mutations as indicated.

able, individually, to direct the fusion proteins to the nucleus (Figure 5).

In a second type of experiment, a sequence of 120 nucleotides encoding the two basic motifs (residues 202–231) was synthesized by polymerase chain reaction and cloned back into pCH110 giving rise to pCHD1+2. In this

construct, the sequence encompassing both motifs was able efficiently to promote nuclear accumulation of β-galactosidase (Figure 5).

In conclusion, using two complementary approaches, we showed that each basic motif is essential but not sufficient on its own to target the chimeric protein into the nucleus.

Correct positioning of a basic residue in the second motif is required independently of the spacer length between the two basic clusters

From the work of Robbins *et al.* (1991) and from the comparison of the published potential bipartite nuclear targeting sequences (Garcia-Bustos *et al.*, 1991; Dingwall and Laskey, 1991; Schreiber *et al.*, 1992) it appears that the spacer region does not have to be a precise length. In order to examine the effect of varying the linear distance between the two basic motifs in terms of structural alteration and/or crucial positioning of the essential K222 amino acid, deletion and insertion experiments were performed.

Modification of the linker region leading to a deletion of two residues in the mutant MD1 had no effect on the targeting function (Figure 6).

Recent NMR studies of a 30 amino acid peptide corresponding to the bipartite PARP NLS (residues 202−231) have shown that the region encompassing the second basic cluster has the propensity to form an α-helix and that this cluster has no well defined orientation towards the first basic one (K.Holland and J.F.Lefèvre, personal communication). In view of these findings we chose to mutate the protein in a number of specific positions in order to gain structural information about the PARP NLS. To test whether a specific orientation between the two clusters is necessary for activity, the YPGD amino acid sequence was inserted between them. This particular sequence was chosen because it would increase the length of the spacer between the two basic clusters and give rise to a YPGDV motif which preferentially adopts a type II β-turn structure as shown by NMR studies in solution (Dyson *et al.*, 1988). Interestingly, this had no effect on protein import (Figure 6), suggesting that no specific orientation between the two basic clusters is required for activity.

Given the findings that the second cluster is part of a nascent α-helical structure as found by NMR (K.Holland and J.F.Lefèvre, personal communication) and that a specific lysine (K222) is required for nuclear localization of the protein, we reasoned that the cytoplasmic phenotype of the mutant K222I could be reverted. To verify this hypothesis, as control experiments we introduced an alanine at position 214 into both the wild-type sequence and the mutant K222I, resulting in mutants MI2 and MI3 respectively. As the insertion occurs in a probably non-structured region, the resulting phenotypes, as expected, were not modified (Figure 6). Moreover, an alanine inserted just before the KKKSKK motif (mutant MI4) did not change the nuclear phenotype. Conversely, the same insertion before the KIKSKK sequence (mutant MI5) partially reverted the cytoplasmic phenotype due to the mutation K222I (Figure 6). This result strongly suggests that a helical-type structure containing a lysine in a critical position is necessary for activity.

Discussion

In this study, we identified a sequence of the human poly(ADP-ribose) polymerase with nuclear targeting potential. The sequence $_{207}$KRK-X(11)-KKKSKK$_{226}$, included in the DNA-binding domain, leads to rapid and total nuclear accumulation when fused to a cytoplasmic protein. To characterize this NLS, two complementary strategies were used: (i) a subtractive approach by which we identified residues which are necessary for nuclear targeting and (ii) an additive approach by which we determined the amino acids that are sufficient for this function. Our results highlight the importance of using these two approaches concomitantly in the study of an NLS sequence. As a reporter protein we chose β-galactosidase, whose activity can be rapidly and easily monitored by histochemical staining (Sanes *et al.*, 1986). However, it is known that β-galactosidase is active as a tetramer while PARP is assumed to form dimers (Bauer *et al.*, 1990). Moreover, the multimerization of a partially defective monomeric signal is sufficient for nuclear targeting, while one monomer is not functional on its own (Roberts *et al.*, 1987). Consequently, a mutation which affects nuclear localization implies that a real alteration in the function of the corresponding signal has occurred.

The cytoplasmic retention resulting from deletion of either of the basic motifs (mutants NPARPΔ1 and NPARPΔ2), means that the PARP NLS is larger than the prototype NLS, that of SV40 large T-antigen (Kalderon *et al.*, 1984). Interestingly, we found that the second motif, $_{221}$KKKSKK$_{226}$, which resembles the solitary karyophilic signal of T-antigen, is not sufficient on its own to direct the fusion protein to the cell nucleus and requires basic amino acids located 11 residues upstream (Figure 1). Our results demonstrate that the PARP NLS is made up of two basic clusters separated by 11 intervening amino acids and belongs to the class of bipartite signals previously described by Robbins *et al.* (1991). Poly(ADP-ribose) polymerase has a modular structure: from limited proteolytic degradation of PARP (Buki and Kun, 1988) we know that the second motif is located at the surface of the protein since this sequence is preferentially cleaved by plasmin. It is interesting to point out that the proteolytic cleavage site (K222−K223) involves the amino acid residue (K222), essential for nuclear targeting, identified by site-directed mutagenesis in the second basic cluster.

It is easy to envisage that the two positively charged motifs could adopt a conformation recognizable by the transport machinery. The structure of the motif responsible for the targeting function of proteins remains to be elucidated. It is tempting to speculate that the conservation of this fundamental function for the cell is not strictly dependent on a well defined sequence and that variability of this sequence (in the spacer region but also in the basic motif) can be tolerated if one basic residue is able to occupy the correct position to interact with the receptor protein(s). This could be the reason why basic amino acid stretches longer than necessary have been selected throughout evolution.

Materials and methods

DNA constructs

pCH constructs. The mammalian expression vector used for the expression of the DNA-binding domain of PARP in these experiments was pCH110 (Pharmacia). This vector allowed the production of a β-galactosidase fusion protein driven by the SV40 early promoter. The 733 bp 5′ cDNA fragment encoding the PARP N-terminal domain was isolated from the plasmid pTG161FWT by digestion with *Hin*dIII (Gradwohl *et al.*, 1990) and inserted into the unique *Kpn*I site using a *Kpn*I−*Hin*dIII adaptor (Haymerle *et al.*, 1986) resulting in pCHNPARP.

For site-specific mutagenesis, the PARP insert from pCHNPARP was

cloned into M13 mp18, from which the single stranded template was prepared. Oligonucleotide-directed mutagenesis was performed essentially according to the Amersham kit protocol. All point mutants were named according to the following nomenclature: the first letter represents the wild-type amino acid followed by its position in the sequence and the second letter indicates the mutated one. The mutants NPARPΔ1 and NPARPΔ2 resulted in the deletion of residues 207–209 and 221–226 respectively.

After mutagenesis, cDNAs were sequenced by the dideoxynucleotide chain termination method (Sanger and Coulson, 1975) to verify that the desired changes were obtained, and cloned back into pCH110.

Vectors expressing the two domains D1 and D2 independently were constructed by insertion at the *Kpn*I site of pCH110 appropriately annealed oligonucleotides. To prepare the vector pCHD1+2, PCR was used to generate a fragment of PARP encoding both basic motifs from valine 202 to asparagine 231.

pCHSV resulted in the cloning of the sequence of the SV40 large T-antigen NLS ($_{125}$PPKKKRKVEDPV$_{136}$; Kalderon *et al.*, 1984) using two complementary oligonucleotides, which were designed to harbour *Kpn*I sites at their extremities after annealing.

pECV constructs. The full-length cDNA or the deleted mutant Δ2 were cloned as *Pst*I inserts into the unique *Xho*I site of the pECV23Xho vector (Belt *et al.*, 1989) using a *Xho*I–*Pst*I adaptor, as pECVPARP and pECVPARPΔ2 respectively.

pTG constructs. The entire coding sequences of PARP and PARPΔ2 cDNAs were cloned as *Pst*I inserts into the *Pst*I site of the prokaryotic expression vector pTG161 (Courtney *et al.*, 1984) giving pTGPARP and pTGPARPΔ2 respectively. Transcription of the cloned sequences proceeds from the major leftward promoter of bacteriophage λ P_L which is controlled by the host encoded thermosensitive repressor cI 857. Overexpression of the cloned sequence was obtained after a temperature shift from 28 to 37°C of the *E.coli* strain TGE900 (Transgene, Strasbourg).

Blotting experiments

Proteins transferred to nitrocellulose sheets were probed for their DNA binding capacity by South-western blotting (Mazen *et al.*, 1989) and their poly(ADP-ribose) polymerase activity was tested as described by Simonin *et al.* (1990, 1991).

Cell culture and DNA transfection

HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 5% fetal calf serum and 20 mM L-glutamine. Cells were transfected with 10 μg of plasmid DNA (pCH or pECV constructs) by the calcium phosphate precipitation method (Graham and van der Eb, 1973). At 48 h following transfection, the localization of the β-galactosidase activity of chimeric proteins expressed by the pCH constructs was revealed by staining with X-Gal. In cells transfected with pECV constructs, expression of the full-length enzyme or the D2 deleted version was visualized by immunofluorescence as described below.

Histochemical staining procedure

At 48 h after transfection, cells were washed with three changes of TBS (50 mM Tris–HCl, pH 7.4, 150 mM NaCl) and fixed for 10 min at 0°C in the same buffer with 2% formaldehyde and 2% glutaraldehyde. The fixed cells were washed twice with TBS and covered with 10 ml of the β-galactosidase reaction mix (5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2 mM $MgCl_2$ and 0.1 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal). The reaction was carried out for 2–4 h at 37°C (Sanes *et al.*, 1986)

Indirect immunofluorescence staining

Transfected cells grown on coverslips were fixed in 100% ethanol at −20°C for 10 min. The coverslips were incubated for 45 min at 37°C with a polyclonal antibody raised against the human PARP second zinc finger peptide (residues 122–165; Simonin *et al.*, 1991). After washing for 15 min in PBS (130 mM NaCl, 20 mM Na_2HPO_4, 10 mM KH_2PO_4) coverslips were incubated with TRITC-conjugated anti-rabbit immunoglobulin (Sigma) for 30 min at 37°C. Antibodies were diluted 100-fold in PBS containig 1% bovine serum albumin (w/v). After washing for 15 min in PBS, immunofluorescence was evaluated using a Leitz Dialux microscope.

Acknowledgements

This work was supported by the Association pour la Recherche contre le Cancer by the Fédération Nationale des Ligues contre le Cancer and by the Ligue Nationale contre le Cancer (Comité départemental du Haut-Rhin). We wish to thank O.Poch, K.Holland, J.Kleinschmidt and Y.Boulanger for helpful comments and discussions.

References

Althaus,F.R. and Richter,C. (1987) *Mol. Biol. Biochem. Biophys.*, **37**, 1–126.

Bauer,P.I., Buki,K.G., Hakam,A. and Kun,E. (1990) *Biochem. J.*, **270**, 17–26.

Belt,P.B.G.M., Groeneveld,H., Teubel,W.J., van de Putte,P. and Backendorf,C. (1989) *Gene*, **84**, 407–417.

Benjamin,R.C and Gill,D.M. (1980) *J. Biol. Chem.*, **255**, 10502–10508.

Buki,K.G. and Kun,E. (1988) *Biochemistry*, **27**, 5990–5995.

Chelsky,D., Ralph,R. and Jonak,G. (1989) *Mol. Cell. Biol.*, **9**, 2487–2492.

Courtney,M., Buchwalder,A., Tessier,L.H., Jaye,M., Benavente,A., Balland,A., Kohli,V., Lathe,R., Tolstoshev,P. and Lecocq,J.P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 669–673.

de Murcia,G., Ménissier-de Murcia,J. and Schreiber,V. (1991) *BioEssays*, **13**, 455–462.

Dingwall,C. (1991) *BioEssays*, **13**, 213–218.

Dingwall,C. and Laskey,R.A. (1986) *Annu. Rev. Cell Biol.*, **2**, 367–390.

Dingwall,C. and Laskey,R.A. (1991) *Trends Biochem. Sci.*, **16**, 478–481.

Dingwall,C., Sharnick,S.V. and Laskey,R.A. (1982) *Cell*, **30**, 449–458.

Dyson,H.R., Rance,M., Houghten,R.A., Wright,P.E. and Lerner,R.A. (1988) *J. Mol. Biol.*, **201**, 161–200.

Feldherr,C.M., Kallenbach,E. and Schultz,N. (1984) *J. Cell Biol.*, **99**, 2216–2222.

Garcia-Bustos,J., Heitman,J. and Hall,M.N. (1991) *Biochim. Biophys. Acta*, **1071**, 83–101.

Gradwohl,G., Ménissier-de Murcia,J., Molinete,M., Simonin,F., Koken,M., Hoeijmakers,J.H.J. and de Murcia,G. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 2990–2994.

Graham,F.L. and van der Eb,A.J. (1973) *Virology*, **52**, 456–467.

Guiochon-Mantel,A., Lescop,P., Christin-Maitre,S., Loosfelt,H., Perrot-Applanat,M. and Milgrom,E. (1991) *EMBO J.*, **10**, 3851–3859.

Hall,C.V., Jacob,P.E., Ringold,G.M. and Lee,F. (1983) *J. Mol. Appl. Genet.*, **2**, 101–109

Haymerle,H., Herz,J., Bressan,G.M., Frank,R. and Standlay,K.K. (1986) *Nucleic Acids Res.*, **14**, 8615–8624.

Huppi,K., Bhatia,K., Siwarski,D., Klinman,D., Cherney,B. and Smulson,M. (1989) *Nucleic Acids Res.*, **17**, 3387–3401.

Ittel,M.E., Garnier,J.M., Jeltsh,J.M. and Niedergang,C.P. (1991) *Gene*, **102**, 157–164.

Kalderon,D., Richardson,W.D., Markham,A.F. and Smith,A.E. (1984) *Nature*, **311**, 33–38.

Kameshita,I., Matsuda,Z., Tanigushi,T. and Shizuta,Y. (1984) *J. Biol. Chem.*, **259**, 4770–4776.

Lanford,R.E. and Butel,J.S. (1984) *Cell*, **37**, 801–813.

Lanford,R.E., Kanda,P. and Kennedy,P. (1986) *Cell*, **46**, 575–582.

Mazen,A., Ménissier-de Murcia,J., Molinete,M., Simonin,F., Gradwohl,G., Poirier,G.G. and de Murcia,G. (1989) *Nucleic Acids Res.*, **17**, 4689–4698.

Ménissier-de Murcia,J., Molinete,M., Gradwohl,G., Simonin,F. and de Murcia,G. (1989) *J. Mol. Biol.*, **210**, 229–233.

Newmeyer,D.D., Lucocq,J.M., Bürglin,T.R. and de Robertis,E.M. (1986) *EMBO J.*, **5**, 501–510.

Rihs,H.P. and Peters,R. (1989) *EMBO J.*, **8**, 1479–1484.

Rihs,H.P., Jans,D.A., Fan,H. and Peters,R. (1991) *EMBO J.*, **10**, 633–639.

Robbins,J., Dilworth,S.M., Laskey,R.A. and Dingwall,C. (1991) *Cell*, **64**, 615–623.

Roberts,B. (1989) *Biochim. Biophys. Acta*, **1008**, 263–280.

Roberts,B.L., Richardson,W.D. and Smith,A.E. (1987) *Cell*, **50**, 465–475.

Saito,I., Hatakeyama,K., Kido,T., Ohkubo,H., Nakanishi,S. and Ueda,K. (1990) *Gene*, **90**, 249–254.

Sanes,J.R., Rubinstein,J.L.R. and Nicolas,J.-F. (1986) *EMBO J.*, **5**, 3133–3142.

Sanger,F. and Coulson,A.R. (1975) *J. Mol. Biol.*, **94**, 441–448.

Schreiber,V., de Murcia,G. and Ménissier de Murcia,J. (1992) *Médecine/sciences*, **8**, 134–139.

Simonin,F., Ménissier-de Murcia,J., Poch,O., Muller,S., Gradwohl,G., Molinete,M., Penning,C., Keith,G. and de Murcia,G. (1990) *J. Biol. Chem.*, **265**, 19249–19256.

Simonin,F., Briand,J.P., Muller,S. and de Murcia,G. (1991) *Anal. Biochem.*, **195**, 226–231.

Stewart,M., Whytock,S. and Mills,A.D. (1990) *J. Mol. Biol.*, **213**, 575–582.

Uchida,K., Morita,T., Sato,T., Ogura,T., Yamashita,R., Nogushi,S., Suzuki,H., Nyunoya,H., Miwa,M. and Sugimura,T. (1987) *Biochem. Biophys. Res. Commun.*, **148**, 617–622.

Received on December 27, 1991; revised on April 15, 1992