

Investigation of Fanconi Anemia Protein Interactions by Yeast Two-Hybrid Analysis

Pia A. J. Huber,*,1 Annette L. Medhurst,* Hagop Youssoufian,† and Christopher G. Mathew* *Division of Medical and Molecular Genetics, Guy's, King's and St. Thomas' School of Medicine, Guy's Hospital, 7th Floor, Guy's Tower, London SE1 9RT, United Kingdom; and †Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas

Received December 7, 1999

Fanconi anemia is a chromosomal breakage disorder with eight complementation groups (A-H), and three genes (FANCA, FANCC, and FANCG) have been identified. Initial investigations of the interaction between FANCA and FANCC, principally by coimmunoprecipitation, have proved controversial. We used the yeast two-hybrid assay to test for interactions of the FANCA, FANCC, and FANCG proteins. No activation of the reporter gene was observed in yeast coexpressing FANCA and FANCC as hybrid proteins, suggesting that FANCA does not directly interact with FANCC. However, a high level of activation was found when FANCA was co-expressed with FANCG, indicating strong, direct interaction between these proteins. Both FANCA and FANCG show weak but consistent interaction with themselves, suggesting that their function may involve dimerisation. The site of interaction of FANCG with FANCA was investigated by analysis of 12 mutant fragments of FANCG. Although both N- and C-terminal fragments did interact, binding to FANCA was drastically reduced, suggesting that more than one region of the FANCG protein is required for proper interaction with FANCA. © 2000 Academic Press

Key Words: Fanconi anemia; bone marrow failure; cancer susceptibility; yeast two-hybrid; chromosomal breakage.

Fanconi anemia (FA) is a rare genetic, autosomal recessive disorder associated with progressive bonemarrow failure, various congenital abnormalities and

Abbreviations used: FA, Fanconi anemia; PMSF, phenylmethylsulfonyl fluoride; GAL4 AD, activation domain of a yeast transcription factor; GAL4 BD, DNA binding domain of a yeast transcription factor; FANCA, Fanconi anemia complementation group A protein (gene); FANCC, Fanconi anemia complementation group C protein (gene); FANCG, Fanconi anemia complementation group G protein (gene); SV40 large T A, SV40 large T antigen.

¹ To whom correspondence should be addressed. Fax: +44 (0) 207 955 4644. E-mail: pia.huber@kcl.ac.uk.

early occurrence of malignancies (1). Cells from FA patients show cell cycle disturbances and are hypersensitive to the clastogenic effect of DNA cross-linking agents. The detection of at least eight complementation groups suggests that mutations or deletions in eight or more genes are responsible for the disease (2, 3). Mutations in the three genes cloned so far, FANCA, FANCC, and FANCG, account for approximately 85% of FA patients (2, 4-6). The proteins show no sequence similarity to each other or to any other protein in GenBank, and their cellular function is as yet unknown. Two basic properties of the gene products have been targeted for investigation to illuminate their molecular function: cellular localisation and determination of interactive partners.

Studies of the localisation of the FA proteins FANCA and FANCC using immunofluorescence, cellular fractionation combined with immunoblotting, and overexpression of GFP fusion proteins have produced conflicting results. The first localisations for FANCC identified the protein mainly in the cytoplasm, (7-11). However, further studies have suggested at least a partial localisation in the nucleus (12), and that translocation of FANCC to the nucleus may require formation of a complex with FANCA (13). The localisation of FANCA posed similar problems, although a bipartite nuclear localisation signal had been identified within its sequence (6). Early reports again found FANCA to be located in the cytoplasm (11), but it is apparent that FANCA also localises to the nucleus (13-15), and two recent detailed immunofluorescence studies observed FANCA mainly in the nucleus (16, 17). Initial studies of FANCG also indicate a predominantly nuclear localisation (18, 19).

Studies of the interaction of the FANCC and FANCA proteins, mainly by co-immunoprecipitation analysis, have again produced conflicting results, with one group reporting interaction between these two proteins (13, 14), and another finding no interaction (10, 15). The reasons for these differences have not yet been resolved



(20). If the interaction is relatively weak or transient, it might be difficult to detect. Alternatively, it could be argued that the interaction occurs only *in vitro*, under specific experimental conditions, or that it requires post-translational modification of one or both proteins. There is evidence that phosphorylation of the FANCA protein is required for this interaction (21). Two groups have recently reported interaction between FANCA and FANCG (18, 19).

The question of which FA proteins interact is central to our understanding of the cellular pathway that leads to chromosomal fragility in FA. The analysis of these interactions has thus far been based primarily on coimmunoprecipitation experiments. One possible means of addressing some inconsistencies observed is to use a different experimental approach to test for interaction. The yeast two-hybrid assay allows testing of suspected interactions between known proteins as well as screening of libraries for interactive partners with a protein of interest (22). In this system, the DNA binding domain (BD) and the transcription activation domain (AD) of a yeast transcription factor are fused separately to different proteins. If these proteins interact with each other, the two domains are brought into close proximity to trigger transcription (23). One of the advantages of the yeast two-hybrid approach is that it can detect relatively weak and transient interactions. We have employed this system to test for interaction between FANCA, FANCC, and FANCG, for possible dimerisation of the FA proteins, and for the determination of domains involved in binding.

MATERIALS AND METHODS

Bacterial and yeast strains. Escherichia coli strain DH5 α was used for amplification and cloning into vector constructs. The yeast (Saccharomyces cerevisiae) strains Y190 and Y187 were used for the two-hybrid assays. Both strains have the reporter gene LacZ, while Y190 additionally has the HIS3 reporter gene for growth on histidine-minus media.

Plasmids. The vectors used for the two-hybrid assay were purchased from Clontech. The pAS2-1 (bait) vector encodes the DNA binding domain of GAL4 upstream of the polylinker site and includes the selection gene for growth on tryptophan-minus media. The pACT2 and pGADT7 (prey) vectors encode the GAL4 activation domain upstream of the polylinker and include the selection gene for growth on leucine-minus media. The full length FANCA cDNA (6) was excised from the pREP4 construct (gift of Hans Joenje) at the restriction sites NarI and DraI. The fragment, which contains the entire coding region, was then ligated into the DNA binding domain vector pAS2.1 at the compatible sites NdeI and SmaI. The ligation sites were verified by DNA sequencing. The FANCC cDNA (2) was amplified by PCR from a pREP4 construct (gift of Manuel Buchwald), with primers that included suitable restriction sites for subcloning into the pAS2-1 and the pACT2 vectors. The full length FANCG cDNA was subcloned into the bait and prey vectors using RT-PCR for initial amplification from total RNA and subsequent PCR with primers containing suitable restriciton sites for vector ligation. The entire FANCC and FANCG sequences and the linker sites were verified by sequencing on an ABI fluorescent sequencer. FANCG fragments (amino acids 1-101, 99-197, 195-313, 305-413, 393-510, 511-622,

1-313, 305-622, 99-313, 195-510, 195-413) were amplified using the Expand TM High Fidelity PCR System (Boehringer Mannheim) with primers containing restriction sites suitable for ligation into pGADT7. The FANCG fragment inserts of the final constructs were verified by sequencing as described above. The FANCG 1-367 mutant fragment construct was produced in the same way as the FANCG wild type construct. Sequence analysis determined it to be a mutant consisting of residues 1-367 of FANCG and 48 random residues at its C-terminus, as a result of a frameshift deletion of 173 bp. For the production of the FANCA-GAL4-AD construct, the full length sequence of FANCA, was PCR amplified and subcloned into pGADT7 followed by sequence verification as described above for the FANCG fragments. The FANCA insert contains one mutation, A1416 → G(S458G, conserved in mouse; H. van de Vrugt, personal communication). The following control plasmids were provided by Clontech: (1) pVA3.1, expressing murine p53-GAL4 BD; (2) pTD1.1, expressing SV40 large T antigen-GAL4 AD; (3) pCL1, expressing the full length wild type GAL4 protein.

Two-hybrid assay. Simultaneous cotransformation as well as sequential transformation of the yeast strain Y190 was performed with bait and prey vectors according to the Clontech standard protocol. In addition, transformation of Y187 with the bait vector and Y190 with the prey vector followed by yeast mating was used as alternative method of introducing two different plasmids into the same host cell. The yeast mating was executed by combining one colony from the bait-vector transformed yeast (Y187) with one colony of the preyvector transformed yeast (Y190) in 1.5 ml YPD medium and the cell suspension incubated for 18 h at 30°C with shaking at 250 rpm. 150 μ l of the suspension was then plated onto agar with selection for cells transformed by both plasmids. The assay for reporter gene activation was performed using the Filter Lift technique. Filter paper (Whatman) was overlaid onto yeast cells. The cells picked up by the filter paper were lysed by immersion in liquid nitrogen and thawing. The filter paper, colony side up, was placed onto a second filter paper that had been soaked in a buffer containing the β -galactosidase substrate 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal). Cells with interacting hybrid proteins and thus activated reporter gene expression were detected by the development of a blue colour. The weaker reporter gene activations, which probably indicate weaker interactions in our two-hybrid tests, were assessed by a liquid β -galactosidase assay. However, using this assay we were consistently unable to determine values much above background for the weaker interactions, whereas the Filter Lift assays showed a clear distinction between weak and no interaction using a cut-off time of blue colour development of 8 h. We therefore derived a semi-quantitative assessment of the strength of protein interactions by converting our observation of the activation of both reporter genes, HIS3 and LacZ. using the equation

reporter gene activation = $N \times C/t \times \ln T$,

where N is the number of blue colonies (1 for <5 colonies, 2 for 6–50 colonies, 3 for 51–500 colonies, 4 for >500 colonies), C is the intensity of blue colour (dark = 2; light = 1), t is the time of colony growth before LacZ expression test (1 week = 1 or 2 weeks = 2), and T is the time (in minutes) of colour development.

The equation was devised in consideration of the various factors contributing to the turnover of substrate by the transcript of the reporter gene LacZ. We thus assessed the speed and efficiency of growth of the yeast cells transformed with both AD and BD GAL4 constructs. The growth time (t) and the number (N) of colonies surviving the His selection and giving rise to a positive β -galactosidase assay were scored. Since each experiment always included tests involving strong and weak interactions as well as controls, variation of transformation efficiency is excluded from playing a significant role in our analysis. The amount of β -galactosidase expression was assessed semi-quantitatively (i.e. by visual inspection)

TABLE 1
Yeast Two-Hybrid Analysis of FANCA, FANCC, and FANCG

pAS2.1 (bait)	pACT2/pGADT7 (prey)	Reporter gene activation (n)
FANCA	FANCC	- (6)
FANCA	FANCG	+++ (6)
FANCA	FANCA	+ (3)
FANCC	FANCC	- (3)
FANCC	FANCG	- (3)
FANCG	FANCC	- (6)
FANCG	FANCG	+ (4)

Note. Each pAS2-1 construct contains the sequence encoding the DNA-binding domain of the yeast transcription factor GAL4 (residues 1–147) fused in frame to each listed gene. Each pACT2 or pGADT7 construct, respectively, contains the sequence encoding the GAL4 activation domain (residues 768–881) fused in frame to each listed gene. The plus signs indicate activation of the *HIS3* and *LacZ* reporter genes in the yeast and are evaluated in Fig. 3. Listed are the numbers (*n*) of experiments for each score.

as the intensity of the blue colour (dark = 2, light = 1). Finally the speed of blue colour development, reflecting the strength of the association of the protein pair as well as the association-dissociation dynamics, is given in T (minutes) with a cut-off time of 8 h. Since the latter number for T varied greatly for the strong versus the weak reporter gene activation, i.e. interactions, we used $\ln T$ in order to be able to represent these interactions on one graph (see Fig. 2). The values calculated from 2–6 experiments were reproducible throughout and thus confirmed the assessment using our equation.

We have consistently tested all our FA fusion proteins for interaction with control fusion proteins, p53, LaminC and SV40 large T antigen. FA protein interactions were only classed as positive when interaction with control proteins were negative. We found that one of our two-hybrid constructs, pACT2-FANCG, produced activation of the reporter genes when co-transformed with any other BD construct. This "self"-activation was however not observed when we used the mating procedure for fusion protein co-expression. Mating results in less sensitive diploid yeast cells. The mating procedure was thus applied in all our interaction tests involving the construct pACT2-FANCG.

Immunoblotting. Yeast extracts were prepared from cell cultures collected at exponential growth phase by centrifugation. The cell pellets were frozen at -70°C and re-suspended in cracking buffer (8 M urea, 5% SDS, 40 mM Tris-HCl, pH 6.8, EDTA, 0.4 mg/ml bromophenol blue, 100 mM mercaptoethanol with the protease inhibitors 6 μ g/ml pepstatin, 3.6 μ g/ml chymostatin, 3.6 μ g/ml soybean trypsin inhibitor, 3.6 μg/ml leupeptin, 9 mM benzamidin, 23 μg/ml aprotinin, 20 mM PMSF). 100 µl cracking buffer per 7.5 OD₆₀₀ units of cells was used. Glass beads were added to break open the cells and the mixture was vortexed vigorously for 1 min, followed by centrifugation at 13,000 rpm for 5 min. The supernatants were saved and the cell pellets with the glass beads extracted once more with cracking buffer including an additional heating step (100°C for 3-5 min). The proteins in the combined supernatants (10-50 μ l) were separated on an 8-18% polyacrylamide gradient gel (20 \times 20 cm). The proteins were then blotted onto nitrocellulose or onto a polyvinylidene difluoride membrane (Amersham) using a semidry or a submerged blotter (Bio-Rad). Membranes were blocked with PBS-T (0.05% Tween 20) with 5% low fat, dried milk (Marvel). Anti-FANCC and anti-FANCA polyclonal antibodies (8, 24) and anti-GAL4 AD (activation domain) monoclonal antibodies (Clontech) were used for probing. Blots were subsequently incubated in horseradish peroxidase labelled secondary antibodies (Amersham). The bands were

detected by enhanced chemiluminescence according to the manufacturer's directions (Amersham).

RESULTS AND DISCUSSION

We have tested for interaction of the known Fanconi anemia proteins FANCA, FANCC, and FANCG using the yeast two-hybrid assay. Although this approach results in forced translocation of test proteins to the nucleus, it has some advantages over assays such as co-immunoprecipitation of *in vitro* expressed proteins or cellular protein extracts, since no manipulation of intracellular conditions is required, and it is able to detect weak or transient interactions. We have expressed fusion proteins consisting of the full length FA proteins, FANCA, FANCC and FANCG linked to the DNA binding domain (BD) and the activation domain (AD) of the yeast transcription factor GAL4. Interaction of the fusion proteins in yeast cells was measured by the activation of two reporter genes, HIS3 and LacZ. Co-expression of two test fusion proteins was carried out both by co-transformation of the yeast strain Y190 and by a mating assay (see Materials and Methods). Our results, each of which is based on 3-6 experiments, are summarised in Table 1. No activation of the

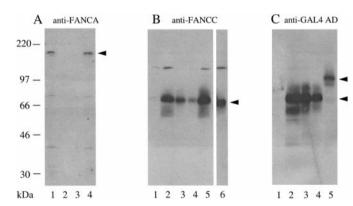


FIG. 1. Immunoblots of yeast protein extracts. Protein extracts were prepared from yeast cells transformed with the constructs named. The hybrid proteins detected are written in bold letters. Arrowheads point to the bands of FANCA-GAL4 AD in (A), FANCC-GAL4 AD and FANCC-GAL4 BD in (B) and (C), and SV40 large T antigen-GAL4 AD in (C). (A) Immunoblot probed with anti-FANCA antibody. Lanes: (1) FANCA-GAL4 BD (185 kDa) + FANCC-GAL4 AD (85 kDa); (2) p53-GAL4 BD (54 kDa) + SV40 large T antigen (96 kDa); (3) no hybrid protein (nontransformed Y190). (4) FANCA-GAL4 BD (185 kDa). (B) Immunoblot probed with anti-FANCC antibody. Lanes: (1) No hybrid protein (nontransformed Y190); (2) FANCA-GAL4 BD (185 kDa) + FANCC-GAL4 AD (85 kDa); (3) FANCC-GAL4 BD (85 kDa) + FANCC-GAL4 AD (85 kDa); (4) p53-GAL4 BD (54 kDa) + **FANCC-GAL4 AD** (85 kDa); (5) FANCG-GAL4 BD + **FANCC-GAL4 AD** (85 kDa); (6) **FANCC-GAL4 BD** (85 kDa). (C) Immunoblot probed with anti-GAL4 AD antibody. Lanes: (1) No hybrid protein (nontransformed Y190); (2) FANCA-GAL4 BD (185 kDa) + FANCC-GAL4 AD (85 kDa); (3) FANCC-GAL4 BD (85 kDa) + FANCC-GAL4 AD (85 kDa); (4) p53-GAL4 BD (54 kDa) + **FANCC-GAL4 AD** (85 kDa); (5) p53-GAL4 BD (54 kDa) + **SV40** large T antigen-GAL4 AD (96 kDa).

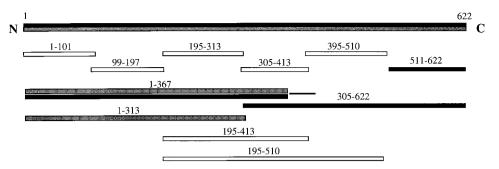


FIG. 2. Peptides of FANCG (full length sequence = amino acids 1–622) involved in its interaction with FANCA and or in dimerisation. Black bars = fragments interacting with full length FANCA; grey bars = fragments interacting with full length FANCG; empty bars = fragments with no interaction observed. The fragment G 1–367 has wild type FANCG residues from 1–367 with additional 48 random amino acids at its C-terminus, indicated by —— (see Materials and Methods).

reporter genes and thus no interaction was found between FANCA and FANCC. In order to check whether this lack of interaction was due to the hybrid proteins not being expressed in our experimental system, we tested protein extracts of the transformed yeast cells by immunoblotting. Figure 1 shows the immunoblots probed with anti-FANCA, anti-FANCC and anti-GAL4 AD antibodies. Bands of the expected size for expression of the appropriate hybrid proteins were detected in each case, confirming that these proteins were expressed in the yeast cells transformed with our constructs. The interaction of FANCA with FANCC (13) is thus not confirmed by our test. This suggests that the interaction may be indirect and mediated by formation of a multimeric protein complex, or that post-translational modification may be required for interaction.

Alternatively, these two proteins may function in separate cellular compartments (15). In contrast, the yeast two hybrid system detected strong interaction of FANCA with FANCG. This confirms the recent reports of interaction of FANCA with FANCG by co-immuno-precipitation (18, 19). We also found that while FANCC seems to exist in monomeric form, as reported previously (9), FANCG as well as FANCA showed weak but consistent interaction with themselves, which suggests that they may be able to form dimers. Dimer formation may be involved in FANCA/FANCG interaction, but this cannot be tested by the yeast two-hybrid assay.

Nevertheless, such a possibility is suggested by our data and will have to be further investigated. We then constructed a series of deletion mutants of the FANCG cDNA in the activation domain vector and tested them for interaction with FANCA in order to map the binding site(s) in FANCG. The fragments generated, and the results of the assays are indicated in Fig. 2. The binding assays are only scored as positive if the interaction was detected consistently over several experiments, and the fragment did not interact with any of the controls (see Materials and Methods). Fragments which contained the N-terminal half or the C-terminal half of FANCG did bind (FANCG₁₋₃₆₇ and FANCG₃₀₅₋₆₂₂),

as did a small C-terminal fragment (FANCG $_{511-622}$). The central region alone does not bind, as shown by the lack of activation with FANCG $_{195-413}$. The interactions between the FANCG fragments and FANCA were very much weaker than with the full length protein. This is illustrated in Fig. 3, which shows a semi-quantitative assessment of binding strength of those fragments which did interact. The drastic reduction in binding strength indicates that at least two different regions of FANCG are required for proper binding. This may reflect co-operative interaction of distinct binding sites, or disruption of the full length protein may result in improper folding, and thus a reduction in binding. Dur-

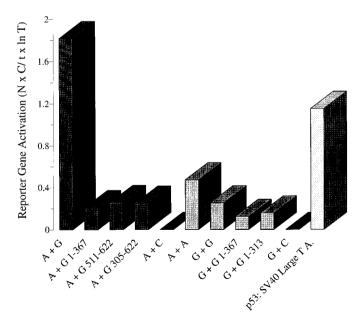


FIG. 3. Intensity of reporter gene activation. The growth of transformed yeast colonies on histidine-lacking media selecting for the activation of the reporter gene, *HIS*, and the Filter Lift assay results, testing for activation of the second reporter gene, i.e., β -galactosidase expression, were evaluated using the formula described under Materials and Methods. Average values are shown, calculated from 2 to 6 experiments.

ing the preparation of this manuscript Kruyt et al. (25) published in vitro co-immunoprecipitation studies showing that two non-contiguous regions in FANCG between amino acids 400-475 and 585-622 contribute to binding to FANCA. Our data are consistent with the involvement of the C-terminal half of FANCG in FANCA binding, and in particular the 111 most C-terminal amino acids. The greatly reduced binding of any of our fragments as well as the fragments tested by Kruyt et al. indicates that co-operative behaviour of various regions of FANCG is involved in binding to FANCA. Our two-hybrid assay detected specific binding additionally in the N-terminal half of FANCG. Indications of co-operative binding of the FANCG molecule contrasts with a defined interaction site on FANCA which was shown to reside in the N-terminal region of residues 1-300 (18), and involves the bipartite NLS in that region (19, 25). However, the C-terminal region of FANCA does appear to be involved in nuclear localisation (16, 17).

In conclusion, our study found no evidence for direct interaction of FANCA and FANCC, suggests that the FANCA and FANCG proteins may form dimers, and provides independent confirmation of direct interaction of FANCA with FANCG. Binding of FANCG to FANCA appears to involve both the amino and carboxy terminal regions of the FANCG molecule. Thus at least some part of the cellular pathway which is disrupted in FA requires co-operative or sequential binding of the products of different FA genes. The yeast two hybrid approach should provide a useful means of investigating interactions between other FA proteins as the genes are cloned, and for detecting novel protein interactions which may illuminate the molecular processes which are disrupted in this disorder.

ACKNOWLEDGMENTS

P.A.J.H. is supported by Cancer Research Campaign, A.L.M. by a studentship from the MRC, UK, and H.Y. by the National Institutes

of Health, HL52138. We thank Neil Morgan and Alex Tipping for practical assistance with construct sequence verification, and Cathryn Lewis for helpful discussions in our statistical evaluation of the yeast two hybrid data.

REFERENCES

- Auerbach, A. D., et al. (1998) in The Genetic Basis of Human Cancer (Kinzler, V., Ed.), pp. 317–331, Mc Graw-Hill.
- Strathdee, C. A., et al. (1992) Nature 358, 763–767 (correction: Nature 358, 434).
- 3. Joenje, H., et al. (1997) Am. J. Hum. Genet. 61, 940-944.
- 4. de Winter, J. P., et al. (1998) Nat. Genet. 20, 281-283.
- The Fanconi Anaemia/Breast Cancer Consortium (1996) Nat. Genet. 14, 324–328.
- 6. Lo Ten Foe, J. R., et al. (1996) Nat. Genet. 14, 320-323.
- Yamashita, T., et al. (1994) Proc. Natl. Acad. Sci. USA 91, 6712-6716.
- Youssoufian, H. (1994) Proc. Natl. Acad. Sci. USA 91, 7975–7979.
- 9. Youssoufian, H., et al. (1995) J. Biol. Chem. 270, 9876-9882.
- 10. Youssoufian, H. (1996) J. Clin. Invest. 97, 2003-2016.
- 11. Kruyt, F. A. E., et al. (1997) Blood 90, 3288-3295.
- 12. Hoatlin, M. E., et al. (1998) Blood 91, 1418-1425.
- 13. Kupfer, G. M., et al. (1997) Nat. Genet. 17, 487-490.
- 14. Näf, D., et al. (1998) Mol. Cell Biol. 18, 5952-5960.
- 15. Kruyt, F. A. E., and Youssoufian, H. (1998) Blood 92, 2229-2236.
- 16. Lightfoot, J., et al. (1999) Hum. Mol. Genet. 8, 1007-1015.
- Walsh, C. E., et al. (1999) Biochem. Biophys. Res. Commun. 259, 594–599.
- Waisfisz, Q., et al. (1999) Proc. Natl. Acad. Sci. USA 96, 10320– 10325.
- 19. Garcia-Higuera, I., et al. (1999) Mol. Cell Biol. 19, 4866-4873.
- Garcia-Higuera, I., and D'Andrea, A. D. (1999) Blood 93, 1430– 1432.
- Yamashita, T., et al. (1998) Proc. Natl. Acad. Sci. USA 95, 13085–13090.
- 22. Fields, S., and Song, O. (1989) Nature 340, 245-246.
- 23. Fields, S., and Sternglanz, R. (1994) Trends Genet. 10, 286-292.
- 24. Kruyt, F. A. E., et al. (1998) Blood 92, 3050-3056.
- 25. Kruyt, F. A., et al. (1999) J. Biol. Chem. **274,** 34212–34218.