

Exon shuffling mimicked in cell culture

(α -crystallin/exon duplication/illegitimate recombination/protein evolution/small heat-shock proteins)

ANKE A. F. VAN RIJK, WILFRIED W. DE JONG, AND HANS BLOEMENDAL[†]

Department of Biochemistry, University of Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

Communicated by Walter Gilbert, Harvard University, Cambridge, MA, May 3, 1999 (received for review October 22, 1998)

ABSTRACT Undesired side products of DNA transfections are usually discarded. However, here, we show that such products may provide insight into mutational events that are also a major driving force in protein evolution. While studying the small heat-shock protein α A-crystallin, we transfected the hamster α A-crystallin gene into a mouse muscle cell line. One of the stable transfected cell lines expressed, in addition to the expected normal α A- and alternatively spliced α A^{ins}-crystallins, two slightly larger, immunologically cross-reacting proteins. These proteins were found to be encoded by a mutant α A-crystallin gene with a large intragenic duplication, arisen by illegitimate recombination at two CCCAT homologies, \approx 1.8 kilobases apart in the normal hamster α A-crystallin gene. As a consequence, a tandem-duplicated exon 3 sequence is present in the mature mRNA of this gene, resulting in a 41-residue repeat in the translated proteins. Cells expressing the elongated α A-crystallins have normal growth characteristics and the usual diffuse cytoplasmic distribution of immunoreactive α A-crystallin. Size-exclusion chromatography of cell extracts indicated that the mutant proteins are readily incorporated into the normal large water-soluble α A-crystallin complexes, showing that the insert does not disturb the integrity of these complexes. This viable α A-crystallin mutant thus mimics the origins and effects of exon duplication, which is a common consequence of exon shuffling in mammalian genome evolution.

It is known that during passage of plasmid DNA in mammalian cells, all sorts of mutations may spontaneously occur in the transfected DNA (1). Therefore, when transfection of a gene into eukaryotic cells yields aberrant products, the logical reaction is to ignore and discard those cells. However, it should be realized that transfection and other recombinant DNA technology simply exploit the mutational processes normally occurring in living organisms. Here, we show that paying attention to “undesired” by-products of transfections may be rewarding in terms of increasing our understanding of mutational and evolutionary mechanisms.

To compare the intracellular properties of the various mammalian small heat-shock proteins (sHsps), we routinely transfect the corresponding genes into different cell lines. Among the six known mammalian sHsps (2), a special position is taken by the α -crystallins (3). α -Crystallin was first discovered as a major structural eye-lens protein. It occurs as large globular complexes of up to 700 kDa, composed of two related types of 20-kDa subunits, α A- and α B-crystallin (4). In rodents and some other mammals, a minor α -crystallin subunit is present, resulting from alternative splicing of the α A-crystallin gene transcript (5–7). This process yields, in addition to the normal 173-residue α A-crystallin, the elongated α A^{ins}-crystallin, with an insertion of 23 amino acid residues. The insert is encoded by the optional exon 2, which is skipped in 80–90% of mRNA because of the presence of an adjacent nonconsensus GC 5' splice site (6, 8). Like other members of

the sHsp family, α -crystallins have *in vitro* chaperone-like activity, suppressing the aggregation of denaturing proteins (9). The most conspicuous *in vivo* feature of α -crystallins is their ability to confer thermotolerance on overexpression in different cell types (10, 11).

Although α B-crystallin occurs in many tissues outside the lens, notably in heart and striated muscle, α A-crystallin is restricted essentially to the lens (12–14). To determine whether α A-crystallin has diverged functionally from α B-crystallin in the course of evolution, it is of interest to establish how α A-crystallin behaves in cell types that are normally the domain of α B-crystallin. Expressing the α A-crystallin gene in a muscle cell line might help to answer this question. Doing so, we observed that one of the stably transfected cell lines expressed, in addition to the expected α A- and α A^{ins}-crystallins, two immunologically cross-reacting proteins at high levels. It seemed informative to identify those proteins and see how viable they are in terms of interfering with normal cellular functioning, complex formation, and chaperone-like capacities. Characterizing these products might shed light on the recombinational processes taking place during passage of the introduced DNA in the mammalian cells, as well as informing us about the properties of spontaneously arising mutants of α A-crystallin. Our analyses identified a mutant α A-crystallin gene with a duplicated exon 3. This finding is in accord with the hypothesis of exon shuffling (15), a theory that has recently been tested experimentally by Gilbert *et al.* (16).

MATERIALS AND METHODS

Plasmid Construction and Transfection. To construct the eukaryotic expression vector for α A-crystallin, an 8.1-kilobase *FspI*–*KpnI* fragment containing the hamster α A-crystallin gene under the vimentin promoter (pVim- α A) was taken from a Puc-18 vector (17). The pVim- α A fragment was ligated into the *XbaI* and *SalI* sites of a pCMV-neo vector, which contains a neomycin-resistance cassette (18). By using the restriction enzymes *XbaI* and *SalI*, the cytomegalovirus (CMV) promoter was deleted from the pCMV-neo plasmid. All ligations were made blunt-end.

Mouse skeletal muscle cells (C2C12) were grown as monolayers in DMEM with glutamax (GIBCO), supplemented with 20% (vol/vol) FCS, penicillin G (100 units·ml⁻¹), and streptomycin (100 ml·ml⁻¹). Cells were transfected based on the method developed by Graham and van der Eb (19) and modified by Wigler *et al.* (20), and stable cell lines were obtained by selection with gentamicin (GIBCO).

SDS/PAGE, Two-Dimensional Electrophoresis, and Immunoblotting. Samples for electrophoresis were prepared by lysing the transfected cells in SDS sample buffer. The protein content of the samples was estimated with the bicinchoninic acid protein reagent (Pierce). One-dimensional SDS/PAGE

Abbreviations: sHsp, small heat-shock protein; pVim- α A, hamster α A-crystallin gene under the vimentin promoter.

[†]To whom reprint requests should be addressed. e-mail: h.bloemendal@bioch.kun.nl.

was performed according to Laemmli's method (21), and two-dimensional gel electrophoresis was performed essentially as described by O'Farrell (22). After one- and two-dimensional gel electrophoresis, Western blotting was performed according to the method of Towbin *et al.* (23). A polyclonal rabbit antiserum against bovine lens α A-crystallin and monoclonal antibody Cr.1-1 (7) against rat α A-crystallin were used.

PCR Amplification and DNA Sequencing. Total RNA was extracted from transfected cells by using Trizol (GIBCO). cDNA was made by 3' rapid amplification of cDNA ends with the aid of a 5'/3' rapid amplification of cDNA ends kit (Boehringer Mannheim). A degenerated forward primer was designed on the basis of a conserved exon 1 sequence in the hamster and mouse α A-crystallin gene (5'-CCA-TTC-AGC-ACC-CTT-GGT-TYA-ARC-G-3') and used in combination with an oligo(dT) and an anchor primer (included in the kit). The PCR product was ligated into the pGEM-T vector (Promega) and sequenced. After sequence determination, the molecular masses and isoelectric points of deduced proteins could be calculated with the aid of the computer program GENERUNNER.

Chromosomal DNA was obtained from transfected cells by using the extraction procedure of Blin and Stafford (24). To amplify the genomic region of interest, a forward primer at the 3' end of exon 3 (5'-GGC-AAG-CAC-AAT-GAG-AGG-CAG-3') was combined with a reverse primer at the 5' end of exon 3 (5'-CTT-GTC-CCG-GTC-AGA-TCG-GAC-3'). The amplification cycle used denaturation at 94°C for 2 min, annealing at 65°C for 1 min, and extension at 72°C for 70 s. The PCR product was again ligated into a pGEM-T vector and sequenced subsequently.

Size-Exclusion Chromatography. The molecular mass of α A-crystallin complexes present in the transfected cell lines was estimated by size-exclusion chromatography. Cells were scraped from culture flasks in lysis buffer, and after centrifugation the soluble fraction was applied onto a Superose 6 HR 10/30 prepac size-exclusion column (Pharmacia). The buffers used for lysis and elution have been described by Lavoie *et al.* (25). Elution was performed at a flow rate of 0.5 ml/min, and 15% of each 1-ml fraction was analyzed by SDS/PAGE and immunoblotting with the monoclonal antibody. Total calf lens protein was used for calibration.

RESULTS

Stable Transfection of the Hamster α A-Crystallin Gene Yields Elongated By-Products. pVim- α A was transfected into

the mouse skeletal muscle cell line C2C12 by using the calcium phosphate precipitation procedure. Among 12 stably transfected lines, 11 showed—to different extents—expression of α A- and α A^{ins}-crystallin, as assayed by Western blotting (not shown). Surprisingly, one of the positive cell lines, pVim- α A7, expressed, in addition to α A- and α A^{ins}-crystallin, two somewhat larger immunologically cross-reacting proteins (Fig. 1A). This cell line and a line with high expression of α A- and α A^{ins}-crystallin (pVim- α A11) were used for further comparison and will be referred to as lines 7 and 11, respectively.

Two-dimensional gel electrophoresis of line 7 cell extract, followed by immunoblotting, showed that the isoelectric points of the two additional proteins lie between those of α A- and α A^{ins}-crystallin (Fig. 1B). Their positions and proportions relative to normal α A- and α A^{ins}-crystallin led us to name them super α A-crystallin and super α A^{ins}-crystallin. From the Western blots in Fig. 1 (see also Fig. 3B), it is difficult to estimate the amounts of the super forms relative to normal α A- and α A^{ins}-crystallin. It is clear, however, that both are expressed at considerable levels. Interestingly, and at variance with earlier α A-crystallin transfections in HeLa cells (26), conspicuous phosphorylation of the expressed proteins seems to occur (arrowheads in Fig. 1B).

As for the viability of the transfected cells, it turned out that line 7 cells, harboring the super forms, grew and divided as readily as cells expressing only the normal α A- and α A^{ins}-crystallins. Also, immunofluorescence studies of cell line 7, as compared with cell line 11, showed no differences in pattern or distribution of α A-crystallin immunoreactivity. In both cell lines, the antiserum gave a similar homogeneous cytoplasmic staining (data not shown).

Sequence Determination Identifies Exon Duplication by Illegitimate Recombination. To establish the nature of the two additional proteins, total RNA was extracted from cell line 7. A 3' rapid amplification of cDNA ends was performed with a forward primer corresponding to nucleotide positions 11–26 in exon 1 of the α A-crystallin gene (see arrowhead a in Fig. 2A), an oligo(dT) reverse primer, and an anchor primer. The amplification products were ligated into a pGEM-T vector and transformed into an *Escherichia coli* DH5 α host strain. Restriction analyses identified four different amplification products. Sequence determination confirmed the presence of normal hamster α A- and α A^{ins}-crystallin mRNAs, whereas the two other products contained a duplicated exon 3 (named 3* in Fig. 2A). No endogenous mouse α A-crystallin sequences were encountered in the products. Translation of the duplicated exon 3 should result in an additional 41 amino acid

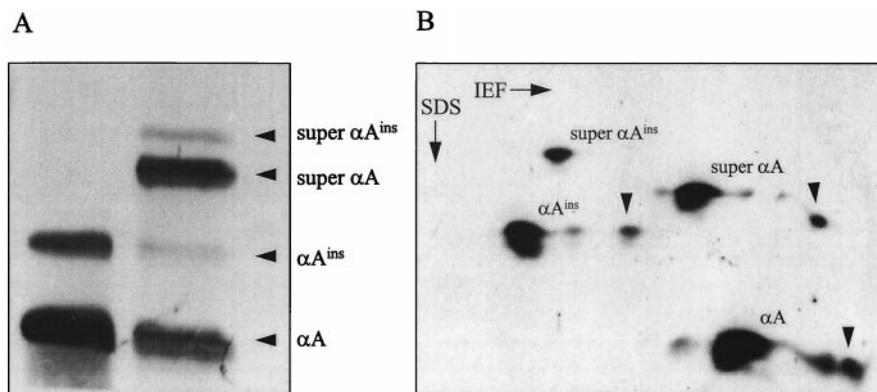


FIG. 1. α A-crystallin immunoreactive proteins present in mouse C2C12 cells after stable transfection with the hamster α A-crystallin gene. (A) Immunoblot after SDS/PAGE of total extracts from transfected cells with a high expression level of normal α A- and α A^{ins}-crystallin with masses of 19.8 and 22.5 kDa, respectively (line pVim- α A11; left lane) and transfected cells expressing two additional α A-crystallin-like proteins: super α A-crystallin and super α A^{ins}-crystallin (line pVim- α A7; right lane). Detection was achieved with a polyclonal anti- α A-crystallin antiserum. (B) Immunoblot after two-dimensional electrophoresis of cell extract from line pVim- α A7. Detection was achieved with a monoclonal anti- α A-crystallin antiserum. Arrowheads indicate presumed monophosphorylation products. The different ratios between the four α A-crystallin-positive proteins in A and B resulted from the use of different antisera and staining conditions.

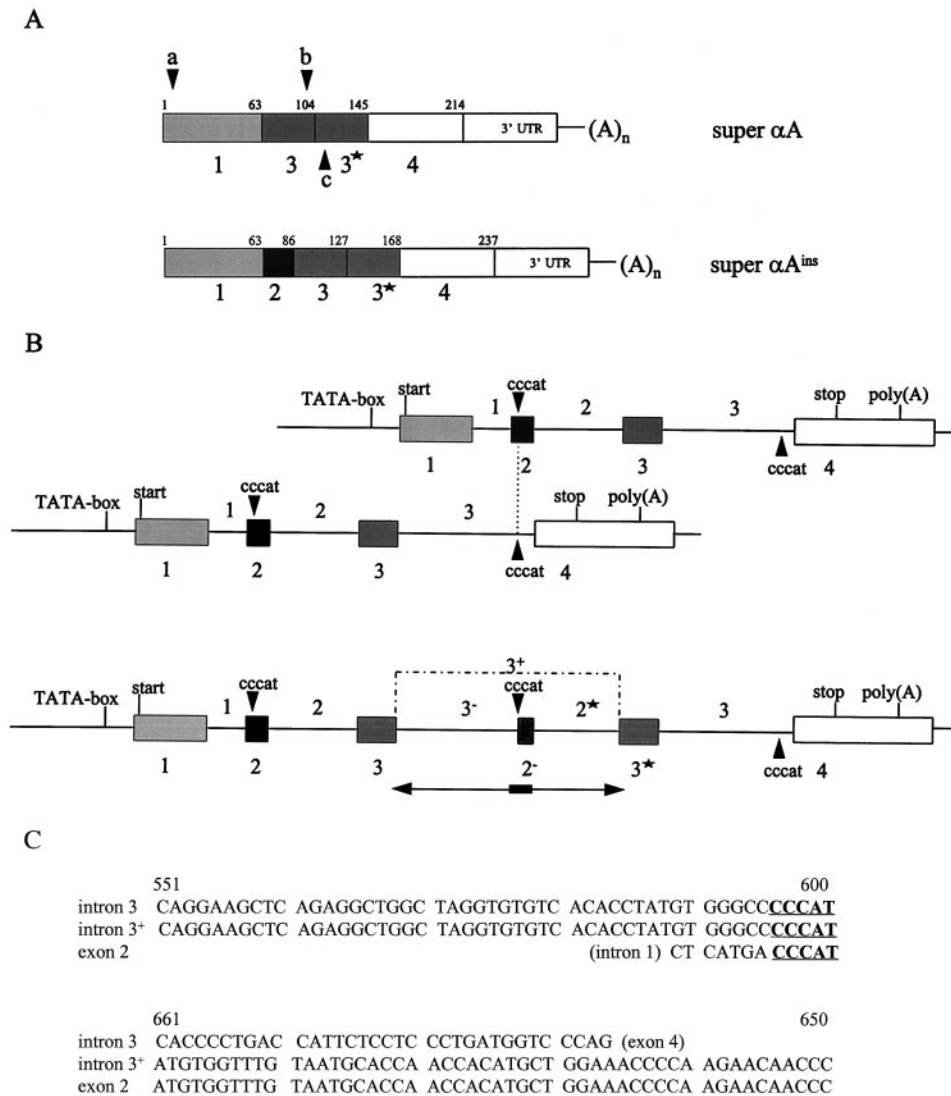


FIG. 2. Exon duplication in the super αA -crystallin gene. (*A*) Schematic representations of the super αA -crystallin mRNAs as deduced from cDNA sequencing. Exons are numbered below the mRNAs. Exon 2 is retained only in the alternatively spliced αA^{ins} -crystallin mRNA. The duplicated exon 3 sequence is indicated as 3*. Numbers above the mRNAs give the amino acid residue positions corresponding to the 3' end of each exon and the total coding region to indicate the lengths of the mutant proteins and inserted regions. Arrowheads point to the positions of the forward primer (a), used to obtain cDNA, and the forward (b) and reverse (c) primers used for sequencing the mutant gene in the genomic DNA. (*B*) Nonhomologous recombination between two copies of the transfected hamster αA -crystallin gene (top and middle genes). Introns are numbered above the genes. Recombination has occurred at the position of the vertical dotted line, apparently facilitated by the CCCAT homologies in exon 2 and intron 3 (arrowheads). As a result, the super αA -crystallin gene (bottom gene) contains a 1.7-kilobase duplication resulting in a truncated intron 3 (indicated as 3⁻) and exon 2 (2⁻), as well as an additional intron 2 (2^{*}) and exon 3 (3^{*}). Because the intron splice site between pseudointron 3⁻ and pseudoexon 2⁻ is lost, these sequences are combined with intron 2^{*} to form the large intron 3⁺ between the duplicated exons 3 and 3*. The double arrow beneath the super gene indicates the region that was amplified by using primers b and c in Fig. 2*A*. The thicker part of this line indicates the position of the sequence presented in Fig. 2*C*. (*C*) The sequence around the illegitimate recombination junction in intron 3⁺ of the super αA -crystallin gene is aligned with the 3' end of intron 3 and the 5' end of exon 2 of the normal hamster αA -crystallin gene. The switch from the intron 3 to the exon 2 sequence occurs at the CCCAT homology at positions 596–600. Numbers above the sequence correspond to the position in the hamster αA -crystallin intron 3 and exon 4.

residues in otherwise normal αA - and αA^{ins} -crystallins. The length and composition of the duplicated sequence correspond well with the observed apparent masses and isoelectric points of the super proteins. The calculated molecular masses are 19.8, 22.5, 24.7, and 27.4 kDa, and the isoelectric points are 6.00, 6.67, 6.02, and 6.52 for normal αA -, αA^{ins} -, super αA -, and super αA^{ins} -crystallin, respectively.

At this stage, it could be concluded that the mutant αA -crystallin gene had originated from nonhomologous or illegitimate recombination between two copies of the transfected hamster gene and that the mutant mRNA was alternatively spliced like the normal rodent αA -crystallin mRNA. Although utterly unlikely in *recA*⁻ *E. coli*, we verified whether the

recombination might have occurred already in the bacteria before DNA isolation and transfection. Thorough PCR analyses of the isolated plasmid gave no indication for such a recombination. To reconstruct how the elongated mutant gene had arisen, we analyzed the structure of this gene in the chromosomal DNA from cell line 7. Because the site of illegitimate recombination should be located between the duplicated exons 3 and 3*, we performed PCR with a forward primer complementary to the 3' end of exon 3 and a reverse primer annealing to the 5' end of exon 3* (arrowheads b and c in Fig. 2*A*). The resulting amplification product was ≈ 1.7 kilobases long. As shown in Fig. 2*B* (see the line under bottom gene), sequence analyses identified, in this order, the expected

5' 21 bp of exon 3; an almost complete intron 3 (3^-), which, at the very end, changed into a 5' truncated exon 2 (2^-); a complete intron 2 (2^*); and finally, the expected 3' 21 bp of the next exon 3 (3^*).

Evidently one hamster α A-crystallin gene has illegitimately recombined at a site located at the 3' end of its intron 3 (Fig. 2*B*, middle gene) with a site at the 5' end of exon 2 of another hamster α A-crystallin gene (Fig. 2*B*, top gene). The sequence of this transition zone in the mutant gene allows us to identify the position of the site of illegitimate recombination more precisely. Up to position 600, this sequence (presented as intron 3^+ in Fig. 2*C*) is identical to the normal intron 3 (top sequence), and from position 596 onward it is identical to the normal exon 2 (bottom sequence). The overlapping 5-bp sequence CCCAT (bold and underlined) is present both in the normal exon 2 and in the normal intron 3 (see arrowheads in Fig. 2*B*) and must have provoked the illegitimate recombination between two copies of the transfected α A-crystallin gene. The recombination event leaves a 5' truncated second exon 2 in the mutant gene (2^- in Fig. 2*B*), which thus becomes part of the third intron of that gene (3^+).

The Super Forms Are Normally Incorporated into the Large α -Crystallin Complexes. One might imagine that a large insertion, such as the 41 amino acid residues encoded by the duplicated exon 3, would interfere with normal folding, stability, and complex formation of α A-crystallin. To assess whether super α A-crystallin and super α A^{ins}-crystallin are indeed incorporated into the large and soluble complexes normally formed by α A-crystallin and α A^{ins}-crystallin, soluble cell extracts from lines 7 and 11 were analyzed by size-exclusion chromatography. After scraping the cells in lysis buffer, the super α -crystallins were found in the soluble fraction, together with the normal forms. The elution profiles of the extracts of cell lines 7 and 11 were very similar (Fig. 3*A*, profiles 2 and 3, respectively). SDS/PAGE of the eluted fractions followed by Western blotting with α A-crystallin antiserum (Fig. 3*B*) showed that, for both cell line 7 and 11, the immunoreactive proteins had elution times similar to control calf lens α -crystallin, between 12 and 30 min (compare profile 1 in Fig. 3*A*). Notably, no later eluting smaller complexes or monomers were observed in either cell line.

The complexes containing super α A-crystallin and super α A^{ins}-crystallin start to elute earlier than the α A-/ α A^{ins}-crystallin complexes (Fig. 3*B*, *Upper vs. Lower*). This fact may be attributed to the mass of the 41-residue insert in the super subunits. Also, the α -crystallin complexes in cell line 7 elute over a broader range than those in line 11, which indicates a greater dispersal of the former complexes. It also seems, as shown in Fig. 3*B* (*Upper*), that the ratios of the four subunits vary with the elution time; the super forms tend to be more prominent in the earlier eluting fractions. However, it may be concluded that even the 64-residue insert in super α A^{ins}-crystallin, as compared with normal 173-residue α A-crystallin, constitutes no detectable hindrance for its regular integration into the large and soluble α -crystallin particles.

DISCUSSION

How Did the Super α A-Crystallin Mutant Arise? Before integration into the genome, transfected circular plasmid vectors are randomly cleaved and may form, by homologous recombination, concatamers in which arrays of monomers are oriented in the same direction (27–29). These concatamers integrate into the chromosomes by illegitimate recombination, doing so preferentially at unstable loci where frequent double-strand breaks and rearrangements occur (28). Joining of transfected and chromosomal DNAs is clearly facilitated by short regions of identity, from 1 to 5 bp, at the junction points of illegitimately recombining DNA molecules (28, 30, 31). Illegitimate recombination requires double-strand breaks, frequently followed by 5' \rightarrow 3' and 3' \rightarrow 5' exonuclease processing, allowing single-strand annealing at short homologous sequences (31). In our case, introduction of the pVim- α A expression vector into mouse C2C12 cells has resulted in illegitimate recombination between two copies of the transfected hamster α A-crystallin gene. The recombination has apparently been triggered by the short homologies CCCAT in exon 2 and intron 3. Interestingly, this sequence also contains a preferred triplet sequence, CAT, for nicking by topoisomerase I, which is known to be directly involved in illegitimate recombination (32). It is difficult to decide whether the illegitimate recombination has occurred before or after the

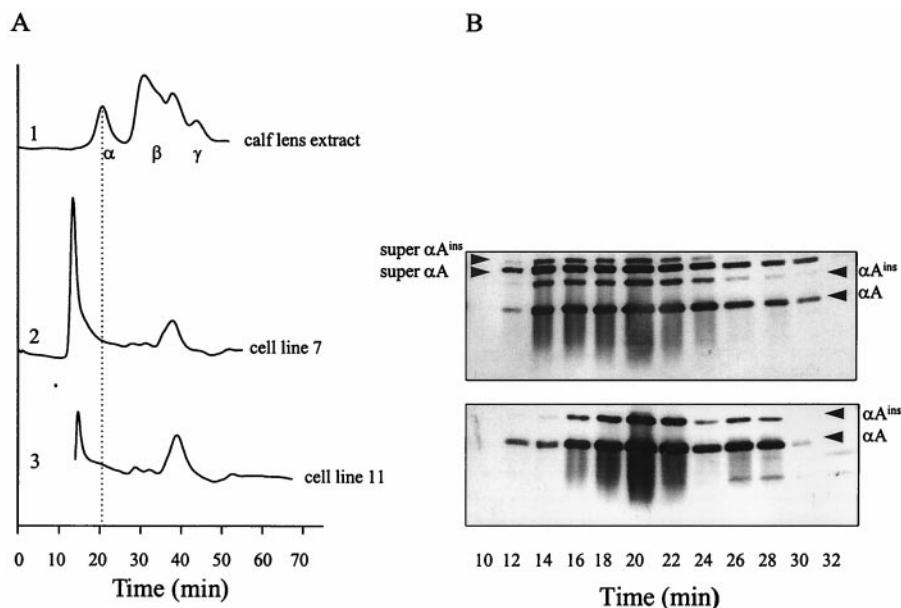


FIG. 3. Complex size of α -crystallin expressed in transfected mouse C2C12 cells. (A) Elution profiles after size-exclusion chromatography (Superose 6) of (i) calf lens soluble proteins, (ii) extract of cell line 7, and (iii) extract of cell line 11. Calf lens α -crystallin complexes have a mass of \approx 700 kDa. (B) Western blotting after SDS/PAGE of Superose 6 fractions from cell line 7 (*Upper*) and cell line 11 (*Lower*). Lanes are numbered according to the elution times shown in A. Blots are stained with a monoclonal anti- α A-crystallin antiserum.

chromosomal integration, although the former seems more likely. In any case, the hamster-specific sequence of the super α A-crystallin gene showed that the endogenous mouse gene had not been involved in the recombination and integration process.

Evolutionary Lessons from a Transfection “Artifact.” Gene transfections make use of the regular DNA repair and rearrangement mechanisms that continuously operate on the mammalian genome. Illegitimate recombination plays a predominant role in these processes (33). However, amazingly little is known about its mechanisms and genetic control (32). Causing deletions, translocations, and amplifications, illegitimate recombination is a rich source of genetic variation, the raw material for evolutionary change. It affects both chromosomal and extrachromosomal DNA. Because extrachromosomal small circular DNAs occur abundantly in eukaryotic cells (30), it is realistic to liken artifacts, such as those described in this paper, to the possible pathways of molecular evolution. Illegitimate recombination is a major innovative force in protein evolution, generating novel combinations of protein domains by exon shuffling (15). Interestingly, although exon shuffling will mostly occur by recombination in introns, our mutant α A-crystallin gene shows that recombination between an exon, albeit a short and optional one, and an intron may occur as well. It has been estimated that at least 19% of all exons have been involved in shuffling events (34). Duplicated exons, too, occur abundantly; at least 6% of all exons in the human genome have recognizably arisen by exon duplication (35). A prerequisite for effective exon duplication is that the involved exons need to be symmetrical with respect to adjacent intron phases. If they are not, the reading frame will be shifted at the junction of the duplicated exons. Also, the viable exon duplication in our super gene has been possible, because the introns flanking exon 3 are both in phase 0, interrupting the reading frame in between two codons. We may conclude that the exon-duplication mutant of α A-crystallin that we have trapped reflects authentic protein evolutionary mechanisms and fits the idea that exon shuffling has a fundamental role in increasing gene diversity (36).

Structural Implications of Exon Duplication in the Super α A-Crystallins. The haphazard insertion, addition, or duplication of a whole domain in a protein that has been molded during hundreds of millions of years of adaptive evolution is most likely to frustrate the integrity of that protein. Only rarely will such a major recombination be favorable, yielding a “hopeful monster” (37) at the protein level, with novel and useful properties as already predicted by Gilbert (15) in his classic 1978 paper. It will most often lead to “hopeless” or at best “neutral monsters”. Our super α A-crystallins probably belong to the latter category, considering that the duplication of 41 residues does not detectably interfere with their *in vivo* viability. The mutant proteins, present at levels similar to the wild-type α A-crystallins (Fig. 1), are not rapidly degraded, have a normal cytoplasmic distribution, do not disturb cell growth, and are incorporated into normal α -crystallin complexes. It seems surprising that a gross structural intrusion can be accommodated by α A-crystallin, which is evolutionarily highly conserved, not having accepted even a single amino acid replacement in the murine rodent lineage for at least 60 million years (38).

This paradox resembles the enigmatic evolutionary origins of α A^{ins}-crystallin, which must have occurred more than 100 million years ago (7). The 23-residue insert in α A^{ins}-crystallin can probably be accommodated as a bulge at the surface or in any internal space of the α -crystallin complex (39). This situation might apply as well for the 41-residue duplication in super α A-crystallin, which is right next to the insert in α A^{ins}-crystallin. Cryo-electron microscopy studies of α B-crystallin complexes identify a highly variable and roughly spherical protein shell around a large central cavity (40). A more regular

hollow spherical complex composed of 24 subunits is reported for the crystal structure of the sHsp from *Methanococcus jannaschii* (MjHSP16.5; ref. 41). It would seem, by comparison, that exon 3 of α A-crystallin codes for three β -strands in the Ig-like fold of the conserved C-terminal domain of the sHsp monomer. The more flexible quaternary structure of the α -crystallins might well allow accommodation of the duplicated β -strands of the super α A-crystallin monomer, either in the central cavity or at the surface. Indeed, an alignment of sHsp sequences from very diverse organisms indicates that major length variation is allowed only in specific regions, notably around the site where the inserts in α A^{ins}- and super α A-crystallins are located (3).

Of course, the survival of a mutant protein in a heterologous cell line is not directly comparable with the “real life” long-term constraints acting in evolution. It nevertheless means that the mutant has passed a first and crucial test, being stable and able to cope with the intracellular environment. Therefore, the super α A-crystallins form a realistic model for studying the evolutionary latitude of the structure–function relationships of the sHsps. Also, in the case of transfection experiments with other genes, it may be rewarding to take advantage of the natural and preselected mutational creativity of the genes themselves, rather than to ignore and discard any unexpected by-products.

We thank Perry Overkamp and Marcel Sweets for technical assistance. This work was supported by European Community-BIOMED (BMH4-CT96-1593).

- Calos, M. P., Lebkowski, J. S. & Botchan, M. R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3015–3019.
- Boelens, W. C., van Boekel, M. A. M. & de Jong, W. W. (1998) *Biochim. Biophys. Acta* **1388**, 513–516.
- de Jong, W. W., Caspers, G. & Leunissen, J. A. M. (1998) *Int. J. Biol. Macromol.* **22**, 151–162.
- Bloemendal, H. (1977) *Science* **197**, 127–138.
- Cohen, L. H., Westerhuis, L. W., de Jong, W. W. & Bloemendal, H. (1978) *Eur. J. Biochem.* **89**, 259–266.
- King, C. R. & Piatigorsky, J. (1983) *Cell* **32**, 707–712.
- Hendriks, W., Sanders, J., de Leij, L., Ramaekers, F., Bloemendal, H. & de Jong, W. W. (1988) *Eur. J. Biochem.* **174**, 133–137.
- Smulders, R. H. P. H., Kokke, B. P. A., Gijzen, M. L. J. & de Jong, W. W. (1998) *Mol. Biol. Rep.* **25**, 225–230.
- Horwitz, J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10449–10453.
- Klemenz, R., Fröhli, E., Steiger, R. H., Schäfer, R. & Aoyama, A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3652–3656.
- van den IJssel, P. R. L. A., Overkamp, P., Knauf, U., Gaestel, M. & de Jong, W. W. (1994) *FEBS Lett.* **335**, 54–56.
- Bath, S. P. & Nagineni, C. N. (1989) *Biochem. Biophys. Res. Commun.* **158**, 319–325.
- Kato, K., Shinohara, H., Kurobe, N., Inaguma, Y., Shimizu, K. & Ohshima, K. (1991) *Biochim. Biophys. Acta* **1074**, 201–208.
- Kato, K., Shinohara, H., Kurobe, N., Goto, S., Inaguma, Y. & Ohshima, K. (1991) *Biochim. Biophys. Acta* **1080**, 173–180.
- Gilbert, W. (1978) *Nature (London)* **271**, 501.
- Gilbert, W., de Souza, S. J. & Long, M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7698–7703.
- Bloemendal, H., Enzlin, J. H., van Rijk, A. F. & Jansen, H. J. (1997) *Exp. Eye Res.* **64**, 1037–1041.
- Bakker, S. J., Markowitz, S., Fearon, E. R., Willson, J. K. V. & Vogelstein, B. (1990) *Science* **249**, 912–914.
- Graham, F. L. & van der Eb, A. J. (1973) *Virology* **52**, 456–467.
- Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. & Axel, R. (1979) *Cell* **16**, 777–785.
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
- O’Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 4007–4021.
- Towbin, H., Staehlin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
- Blin, N. & Stafford, D. W. (1976) *Nucleic Acid Res.* **3**, 2303–2308.

25. Lavoie, J. N., Lambert, H., Hickey, E., Weber, L. A. & Landry, J. (1995) *Mol. Cell. Biol.* **15**, 505–516.
26. van den IJssel, P. R. L. A., Overkamp, P., Bloemendal, H. & de Jong, W. W. (1998) *Biochem. Biophys. Res. Commun.* **247**, 518–523.
27. Coffin, J. M. (1990) *J. Med. Virol.* **31**, 43–49.
28. Merrihew, R. V., Marburger, K., Pennington, S. L., Roth, D. B. & Wilson, J. H. (1996) *Mol. Cell. Biol.* **16**, 10–18.
29. Bishop, J. O. (1996) *Reprod. Nutr. Dev.* **36**, 607–618.
30. Iwasaki, T., Ohki, R., Kiyama, R. & Oishi, M. (1995) *FEBS Lett.* **363**, 239–245.
31. Henderson, G. & Simons, J. P. (1997) *Mol. Cell. Biol.* **17**, 3779–3785.
32. Zhu, J. & Schiestl, R. H. (1996) *Mol. Cell. Biol.* **16**, 1805–1812.
33. Roth, D. & Wilson, J. (1988) in *Genetic Recombination*, eds. Kucherlapati, R. & Smith, G. (Am. Soc. Microbiol., Washington, DC), pp. 621–653.
34. Long, M., Rosenberg, C. & Gilbert, W. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 12495–12499.
35. Fedorov, A., Fedorova, L., Starshenko, V., Filatov, V. & Grigor'ev, E. (1998) *J. Mol. Evol.* **46**, 263–271.
36. de Souza, S. J., Long, M. & Gilbert, W. (1996) *Genes Cells* **16**, 493–505.
37. Goldschmidt, R. (1940) in *The Material Basis of Evolution* (Yale Univ. Press, New Haven, CT).
38. Hendriks, W., Leunissen, J., Nevo, E., Bloemendal, H. & de Jong, W. W. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5320–5324.
39. Smulders, R. H. P. H., van Geel, I. G., Gerards, W. L. H., Bloemendal, H. & de Jong, W. W. (1995) *J. Biol. Chem.* **270**, 13916–13924.
40. Haley, D. A., Horwitz, J. & Stewart, P. L. (1998) *J. Mol. Biol.* **277**, 27–35.
41. Kim, K. K., Kim, R. & Kim, S. (1998) *Nature (London)* **394**, 595–599.