Journal of Chromatography, 615 (1993) 23–36
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 6781

Use of automated smoothing and deconvolution procedures for the determination of human breast cancer estrogen and progesterone receptor isoforms, after high-performance size-exclusion chromatography

H. Cren, C. Lechevrel, G. Roussel, E. Lemoisson and J. Goussard*

Laboratoire d'Analyses Isotopiques, Centre François Baclesse, Route de Lion sur Mer, 14021 Caen (France)

(First received April 27th, 1992; revised manuscript received February 2nd, 1993)

ABSTRACT

Automated smoothing and deconvolution procedures were used to analyse complex chromatographic patterns of human breast cancer estrogen and progesterone receptors, obtained by size-exclusion chromatography. By injecting, first, different known amounts of a radio-iodinated protein in a TSK G-3000 SW column, constructed complex chromatograms (twelve chromatograms) with known peak positions were obtained, and were further treated by mathemetical functions to determine a smooth-deconvolution strategy, which could be used with unknown chromatographic patterns. The determination of peak areas by a "curve-fit" program showed a good correlation with the amounts of protein injected (r = 0.91). Human breast cancer estrogen receptors (56 cases) and progesterone receptors (45 cases) were chromatographed in a TSK G-3000 SW column, and further analysed with the smoothing and deconvolution procedures: eight different estrogen receptor isoforms were detected with molecular masses ranging from 530 000 [Stoke's radius (R_s) = 7.7 nm] to 23 000 (R_s = 2.6 nm), and eight progesterone receptor isoforms were observed with masses ranging from 680 000 (R_s = 8.6 nm) to 50 000 (R_s = 3.1 nm). The dissociation effect of KCl (0 to 1 *M*) on receptor structure yielded different proportions of receptor isoforms, but did not modify their different peak positions.

INTRODUCTION

For twenty years, the ligand binding capacity of estrogen receptors (ER) has been used for the determination of ER in human breast tumours. Clinical investigations showed that ER positivity was a good prognostic factor, specially for patients undergoing endocrine therapy [1]: 70% of ER+ patients (ER > 10 fmol/mg protein) responded to an endocrine therapy, but only 5–10% of ER− patients. Despite the benefit of hormone therapy for ER+ patients, 30% of them did not respond to the endocrine treatment, and qualitative studies were suggested to determine possible modifications in the receptor structure, leading to its inactivation. By sucrose density gradient sedimentation, Wittliff [2] observed that ER from human breast tumours sedimented as two components [8–9S, MW > 300 000 (300 kDa), polymeric form; 4–5S, MW = 70 kDa, monomeric form], and that the proportion of these two forms varied among tumours. The presence of the 8S form, especially, was related to a good response to hormone therapy. In addition to the 8S and 4S ER forms, other structural forms were identified, including fragmented forms of the 4S (3–4S, MW 40–70 kDa), and the meroreceptor (2–3S, MW 20–25 kDa) [3]: it was suggested that these

* Corresponding author.

0378-4347/93/$06.00

different forms were due to proteolytic digestion of ER [4].

High-performance size-exclusion chromatography (HPSEC) offers improved resolution and high reproducibility, in addition to short elution times. With this method, different receptor isoforms can be observed with molecular masses ranging from 23 to 530 kDa, but the complex chromatograms obtained need further mathematical treatment to analyse the overlapping peaks, and to detect in a given breast tumour the presence of particular isoforms that could be eventually responsible for receptor inactivation.

The aim of this work was to develop an analytical method for the rapid and sensitive detection of human breast cancer ER and progesterone receptor (PR) isoforms obtained by HPSEC under different conditions. The smoothing and deconvolution procedures were first studied with constructed known chromatographic patterns, and further used with complex human breast cancer ER and PR chromatograms to determine the different receptor isoforms obtained from mild dissociation of the polymeric receptor form by potassium chloride.

EXPERIMENTAL

Reagents

[2,4,6,7-^{3}H]Estradiol (95 Ci/mmol) was obtained from Amersham International (Little Chalfont, UK). The following chemicals were used: dihydrotestosterone (DHT), diethylstilbestrol (DES), tris(hydroxymethyl)aminomethane, sodium dihydrogenphosphate dihydrate, disodium ethylenediaminetetraacetic acid (EDTA), dithiothreitol (DTT), sodium molybdate, potassium chloride, glycerol (Merck, Darmstadt, Germany); charcoal Norit A, Dextran T-70, ribonuclease, α-chymotrypsin, trypsin inhibitor, ovalbumin, bovine serum albumin, aldolase, human γ-globulin, equine ferritin and human thyroglobulin (Sigma, St. Louis, MO, USA).

Tumours and cytosols

Tissue samples (30–500 mg) from breast cancer specimens were obtained immediately after surgery and kept frozen in liquid nitrogen until use. Tumour tissue was thawed, minced and then homogenized with six volumes (w/v) of TEDG buffer (10 m*M* Tris–HCl, pH 7.4, 1.5 m*M* EDTA, 0.5 m*M* dithiothreitol, 10 m*M* sodium molybdate and 10% v/v glycerol) with a tissue homogenizer Ultra-Turrax (Janke and Kunkel, Staufen, Germany) for five 5-s bursts. All procedures were at 0–2°C. The resulting homogenate was centrifuged at 105 000 *g* for 60 min. The floating lipid layer was removed and the supernatant (cytosol) was collected. Cytosols were stored in liquid nitrogen until use.

Radioligand assay

Microtitration plates were used and all steps were performed at 0–2°C [5]. Each aliquot of cytosol (100 μl) was incubated for 16–18 h at 0°C with 5 n*M* [^{3}H]estradiol in the presence or absence of a 500-fold excess of DES as competitor, and with DHT (100 n*M*) to prevent binding between estradiol and sex binding globulin [6]. Unbound steroid was removed by incubation with an equal volume of a 0.5% suspension of dextran-coated charcoal (0.5% Norit supra A, 0.05% Dextran T-70 in Tris buffer) for 30 min at 0–2°C with continuous mild shaking, and then by centrifugation at 1000 *g* for 10 min. The supernatant, which contained protein-bound [^{3}H]estradiol, was counted for radioactivity. Specific binding was calculated as the difference between total and non-specific binding (NSB). Cytosol protein concentrations were determined by the method of Lowry *et al.* [7], using bovine serum albumin as standard.

High-performance size-exclusion chromatography

All chromatographic separations were performed at 0–2°C with an LKB liquid chromatograph using a Model 2150 solvent pump and a Rheodyne 7125 injection valve [8]. All buffers and cytosols were filtered through a 0.2-μm filter before use. Size-exclusion chromatography was performed on a Spherogel TSK G-3000 SW column (600 mm × 7.5 mm I.D., LKB, Uppsala, Sweden) fitted with a TSK precolumn. Sample application was through a 500-μl sample loop.

Cytosols labelled in the absence of radioinert DES as competitor (total binding) were cleared of free steroid, and aliquots (300–500 μl) of these receptor preparations were injected into the column. For some experiments, aliquots of receptor preparations were treated for a short time (1–3 h) before injection with KCl (0.4 *M*) to allow dissociation of receptor isoforms. The contribution of NSB in the aliquots injected was measured by chromatography of different NSB, and it was always lower than 1% of the total binding, and equally distributed throughout the chromatogram.

Isocratic elution was with PEDG-K buffer (25 m*M* sodium phosphate, pH 7.4, containing 1.5 m*M* Na_2EDTA, 0.5 m*M* DTT, 10 m*M* sodium molybdate and 10% v/v glycerol) containing 0.4 *M* KCl. Elution was carried out at a flow-rate of 0.5 ml/min. The radioactivity of the eluate was directly recorded with a radioisotope detector fitted to the column outlet. The TSK G-3000 SW column calibration was performed with PEDG-K as elution buffer, and standard proteins used were: ribonuclease (MW 13.7 kDa, R_s 1.6 nm); α-chymotrypsin (MW 21.6 kDa, R_s 2.0 nm); trypsin inhibitor (MW 24 kDa, R_s 3.1 nm); ovalbumin (MW 43 kDa, R_s 3.0 nm); bovine serum albumin (MW 68 kDa, R_s 3.5 nm); aldolase (MW 158 kDa, R_s 4.8 nm); γ-globulin (MW 158 kDa, R_s 4.5 nm); ferritin (MW 440 kDA, R_s 6.1 nm); thyroglobulin (MW 669 kDa, R_s 8.5 nm). The void volume (V_0) and total liquid volume (V_t) were determined with Dextran Blue and tritiated water, respectively. The elution volume of the standard proteins and of Dextran Blue were determined spectrophotometrically, and the elution volume of tritiated water was determined with the radioisotope detector. The MW and Stoke's radius of ER and PR isoforms were estimated using linear regression analysis of $(K_{av})^{1/3}$ *versus* R_s or MW of the standards: $K_{av} = (V_e - V_0)/(V_t - V_0)$, where V_e is the elution volume of the protein being studied.

The expression "isoform" is used here to refer to different receptor forms sharing the common property to bind with high affinity their own specific ligand (estrogen for ER, progesterone for PR), but having different MW owing to dissociation of the original low-salt polymeric form, to different forms (isoforms) obtained after mild dissociation at high ionic strength (0.4 *M* KCl).

Radioactivity measurement with the radioisotope detector, and processing of chromatographic data

The radioactivity detection system (Flo-One, Flotec SA, La Queue lez Yvelines, France) consisted of a detector head assembly, and a software package for an IBM compatible computer, presenting a totally automated radioactivity detection system for liquid chromatography. The eluate (0.5 ml/min) entering the system is mixed in a mixing chamber with a scintillation fluid (Ecoscint-H, National Diagnostics, Manville, NJ, USA) supplied by a pump (4 ml/min). The resultant homogeneous liquid is pumped through a flow-cell (2.5 ml) positioned between two photomultiplier tubes (PMTs). The radioactive decay events occurring in the flow-cell are monitored by the PMTs and are continuously registered as raw counts. A selectable range of counting update time (2–60 s) can be selected for data processing (a shorter update time means better resolution and poorer counting). Optimal counting of radioactivity, for the visualization of receptor isoforms, was obtained by using an update time of 20 s (= 1 raw count time), a 2.5-ml flow-cell, and a chromatographic flow-rate of 0.5 ml/min. The time-course of each chromatogram was 60 min and, thus, included the elution time of the total liquid volume (57 min). The accumulated raw counts constituted one run (180 raw counts of 20 s = 180 fractions = 60 min) and were saved as a data file on the hard disk of the computer. They were further mathematically treated by a computer program (Spectra Calc, Galactic Industries, Salem, NH, USA) for presentation of the radioactivity data (Fig. 1).

This program was designed to process chromatographic data with the Fourier function transformation (FFT) and with the Savitzky–Golay (SG) smoothing function [9], in addition to spectral deconvolution and curve-fitting [10].

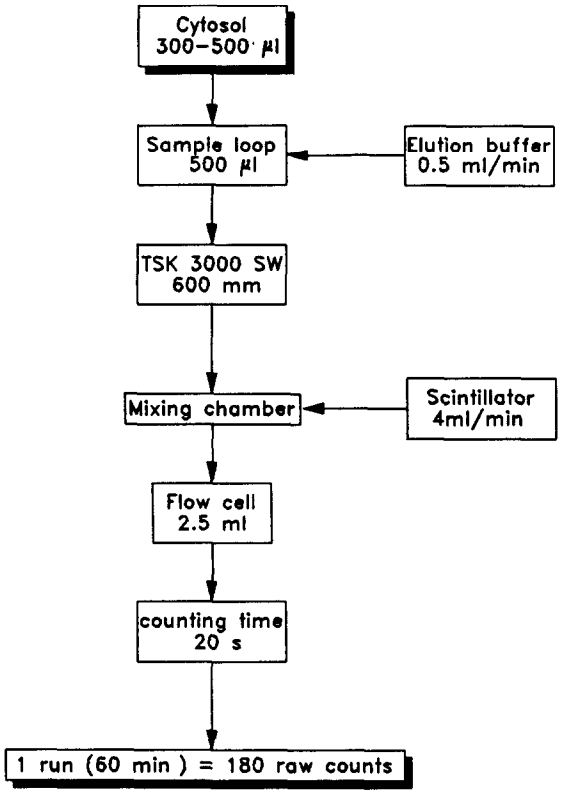

Fig. 1. Chromatographic strategy: the fixed parameters described in Experimental and used to obtain an optimal elution pattern of breast cancer receptor isoforms.

RESULTS

Analysis of single chromatographic peaks

In a first series of experiments, the elution profile of standard proteins used for the calibration of the TSK column were studied. The raw data of the spectrophotometric patterns (Fig. 2A) were imported into the data file of the computer software: a curve-fit program automatically calculated the pure Gaussian band that best fitted the original curve. The program works first by asking the user to make a "guess" at the number of bands present, their peak position, their peak width, and whether they are Gaussian or Lorentzian. The program takes these starting guesses and iterates to find the combination of band heights, positions, and widths that best fit the data file (Fig. 2B). The peak width was determined for each calibration protein, and the measured width average (3.1 ± 0.6 min, ten protein chromatograms) was used as reference width for the width guess asked by the curve-fit program, during the treatment of all complex chromatographic curves composed of overlapping unknown peaks.

The curve-fit program was tested on a series of complex chromatographic curves obtained by successive injections of known amounts of radioiodinated γ-globulin on the TSK column. In this case, the major difficulty encountered during the curve-fit process was the estimation of the centre of each peak. Consequently, the raw data alone were not appropriate to the Gaussian treatment of overlapping chromatographic patterns, such as HPSEC patterns obtained during receptor isoforms analysis. Furthermore, for small peaks, a low signal-to-noise ratio increased the difficulty of this Gaussian treatment. Thus, it was necessary to use further new mathematical functions with raw data, such as smoothing and deconvolution, to improve peak detection.

Comparison of FFT and SG smoothing: application to a Gaussian chromatographic peak

The simple averaging of a large number of runs of the same product enhances the signal-to-noise ratio. Nevertheless, this method consumes biological material and is not convenient for chromatographic analysis of small samples of breast cancer cytosol (300–500 µl); direct computer smoothing operations of chromatograms are more suitable. By acting as a filter, the computer program smooths the noise fluctuations with the aim of avoiding distortions in the recorded data. As the problem of distortion is difficult to assess, we first studied the two smoothing functions, FFT and SG function [9], proposed by the program. These functions were tested on the Gaussian curve obtained by the curve-fit treatment of Fig. 2C and D. The new curve obtained after a first smoothing was treated once again by curve-fit, and the estimated values for the centre, height, width, and area were obtained (Table I). The two smoothing functions, FFT and SG, did not displace the peak centre at any level of smoothing. In contrast, the peak height was

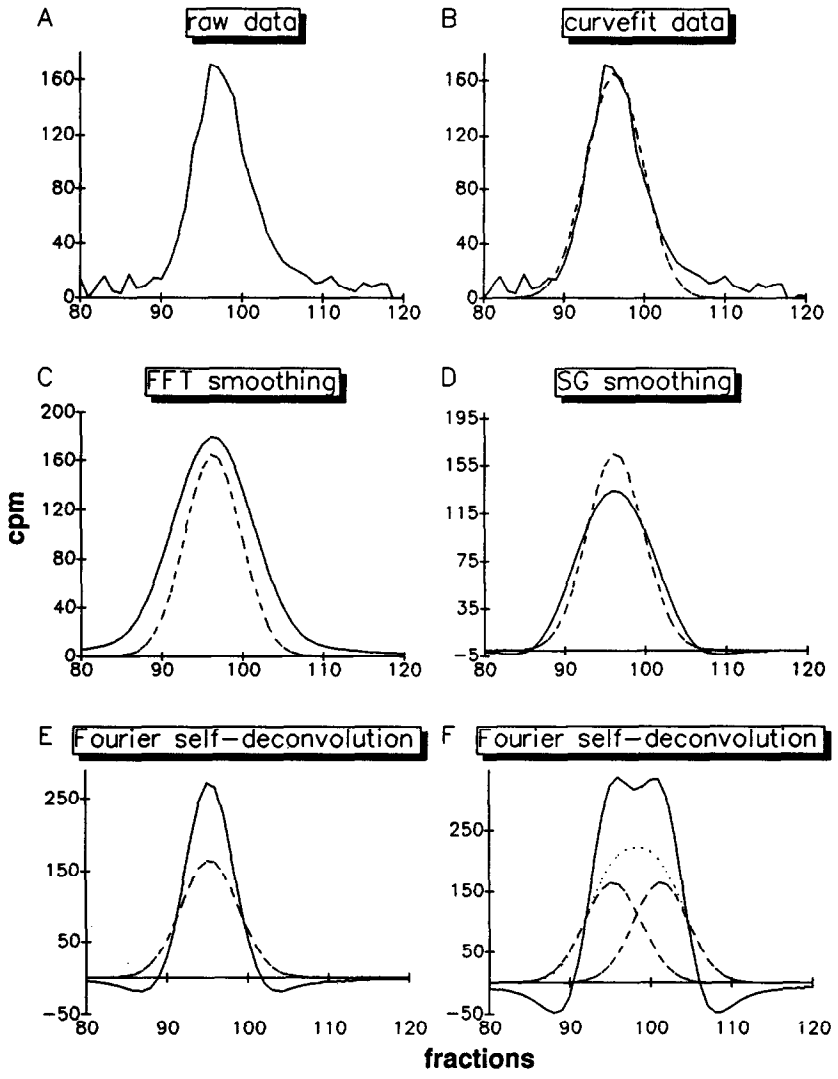

Fig. 2. Analysis of an HPSEC pattern of radiolabelled human γ-globulin using curve-fitting, smoothing, and deconvolution. (A) The chromatographic conditions were as described in Experimental, the activity of the eluate was registered with the Flo-One radioisotope detector, as for receptor isoforms. The registered raw counts were saved as a file on the hard disk of the computer. (B) This file was treated with the curve-fit program. A Gaussian curve was obtained (– – –) which best fitted the chromatographic curve (———), and parameters of this curve were measured: centre (96) and width (10), in counting fractions; height (165) and area (1439) in raw counts. Peak width was defined as the width at one-third of the peak height. (C) and (D) The Gaussian curve (– – –) was smoothed (———) either by the FFT operation (C) or by the SG operation (D), using F_9 and SG_{21} conditions, respectively (see Table I). (E) The Gaussian curve (– – –) was treated with the Fourier self-deconvolution program (———). The deconvolution process (deconvolution constant = 1) reduced the width while increasing the height of the peak. The centre of the Gaussian curve was not displaced by the deconvolution process. (F) Two identical Gaussian curves (– – –) separated by six counting fractions were summed. This new curve (·····) was treated by the Fourier self-deconvolution function (deconvolution constant = 1). The deconvoluted curve (———) showed a double peak, the centres of which corresponded to the centres of the two original Gaussian peaks.

increased with the FFT operation up to 109%, and decreased with the SG operation down to 78%, depending on the smoothing level chosen. The FFT smoothing more rapidly increased the peak width and area (up to 220 and 255%, respectively) than did the SG smoothing (up to

TABLE I

EFFECTS OF THE SMOOTHING LEVEL, OBTAINED EITHER BY THE FOURIER FUNCTION (FFT) OR BY THE SAVITZKY–GOLAY FUNCTION (SG) ON THE VARIATION OF CENTRE, HEIGHT, WIDTH AND AREA OF A GAUSSIAN CURVE

A Gaussian curve obtained by the curve-fit treatment of a pure protein (see Fig. 2B) was smoothed by the FFT or SG smoothing functions. The user can choose among different smoothing levels: F_1 to F_9 for the FFT function, and SG_5 to SG_{21} for the SG function, Results are expressed as a percentage of the ratio of smoothed curve parameters to non-smoothed Gaussian curve parameters.

Smoothing level		Centre		Height (%)		Width (%)		Area (%)	
		FFT	SG	FFT	SG	FFT	SG	FFT	SG
F_1	SG_5	100	100	103	100	100	100	110	102
F_2	SG_7	100	100	103	99	100	100	111	104
F_3	SG_9	100	100	103	98	100	100	113	105
F_4	SG_{11}	100	100	104	96	110	100	115	107
F_5	SG_{13}	100	100	105	94	110	100	119	109
F_6	SG_{15}	100	100	106	90	110	110	124	111
F_7	SG_{17}	100	100	107	86	120	110	135	114
F_8	SG_{19}	100	100	108	82	140	120	159	116
F_9	SG_{21}	100	100	109	78	220	130	255	119

130 and 119%, respectively). Because the most important parameters for qualitative analysis of overlapping peaks were the peak centre and width, we chose to use the SG function for further investigations because this method introduced smaller width distortion than the FFT method. Thus, as the resolution of a peak can be changed by varying the chromatographic conditions (flow-rate, elution buffer, column size, flow-cell of the radioisotope detector, etc.), this resolution can also be changed according to the mathematical functions used. Smoothing function decreased the resolution but reduced the noise. These simultaneous effects complicated the curve-fit treatment of complex chromatographic patterns by fusing overlapping peaks.

Resolution enhancement of overlapping chromatographic peaks by Fourier self deconvolution (FSD)

Complex chromatographic profiles can be considered to consist of several overlapping peaks, each one representing the elution of one component. As SG smoothing reduced both the noise and the resolution of a Gaussian peak or of a chromatographic profile, the effect of smoothing on the peak broadening was corrected by the FSD method [10]. The deconvolution program used two filter parameters: a deconvolution filter constant, and a smoothing filter constant. The smoothing filter constant was maintained at 1 for all treated curves, because it acted as a smoothing function and thus could interfere with the SG smoothing operation. For the deconvolution filter constant, values from 0.2 to 2 could be chosen. The deconvolution process obtained with FSD gave interesting characteristics despite some distortion of the original curve. First, it reduced the peak width and, consequently, enhanced its resolution (Fig. 2E). Second, the peak centre was not modified by the computer treatment. The same results were obtained with two overlapping chromatographic Gaussian peaks (Fig. 2F). The Fourier self-deconvolution failed to separate two Gaussian peaks if they were too much overlapping. A minimum of six fractions between two Gaussian peaks was necessary to separate two peaks by the deconvolution function. The detection of two overlapping proteins by deconvolution also needed six fractions. In other words, a minimum of 2 min (six 20-s counting fractions)

between two overlapping peaks was necessary for their separation by the Fourier self-deconvolution function. The deconvolution studies presented in Fig. 2 were performed with pure Gaussian curves without noise. Complete investigations with the addition of noise in the Gaussian curve were necessary to use further the Fourier self-deconvolution function with complex receptor chromatograms.

Repeated smoothing-deconvolution process of a Gaussian curve to which noise has been added

The noise interferes with chromatographic results mainly if the signal-to-noise ratio is too low. In our studies, the noise had two origins: the radioisotope detector and a residual radioactive background eluted from the column. The smoothing function reduced the noise, but the Fourier self-deconvolution function increased it. These two functions cannot be used with confidence if the user ignores some special disadvantages on the signal-to-noise ratio. For example, false peaks specially due to noise can be enhanced and observed after an overdeconvolution, and sometimes it can be difficult to discriminate between real and false peaks when the signal-to-noise ratio is very low. To avoid such a disadvantage, a strategy was developed including iterated smoothing and deconvolution operations, in addition to the introduction of a new factor that we have called the "resolution ratio" (r.r.). The purpose of this was to prevent the detection of false peaks due to noise. To study the effect of noise in the deconvolution process, the function "add noise" of the program was used with a Gaussian curve. This function allowed the addition of synthetic randomly distributed noise to a data file. As an example, 40% of raw counts of the Gaussian curve of Fig. 2A were added as noise to this curve (Fig. 3A). These worse conditions were never encountered with receptor chromatographic patterns, but they were chosen to determine an r.r. allowing the true discrimination between real and false peaks, without the risk of errors. After noise was added, the curve was smoothed by the SG function with three different smoothing degrees (SG_5, SG_7, SG_9) The curve was further deconvoluted, and the smoothing-deconvolution operations were repeated several times. For each pass, the maximum value of peak counts (P) and of noise counts (N) were registered (N = mean noise + 3 S.D.). A diagram of the variations of P *versus* N is presented in Fig. 3B. This diagram shows that P is related to N ($P = aN^b$; for SG_7: $P = 23.2N^{(0.623)}$), and that the P/N ratio increases with the degree of smoothing, from SG_5 to SG_9. If the curve was not smoothed enough (SG_5), false peaks due to noise were observed in addition to the real peak analysed (arrows, Fig. 3C). In contrast, only one peak was obtained when the SG_7 or SG_9 smoothing degree was used (Fig. 3D). Thus, in this worse noise condition (40% of noise), real peaks were observed only if $P > aN^b$ with $a \leqslant 23.2$, and $b \geqslant 0.623$. From these data, we defined empirically an r.r. $= P/aN^b$, the value of which should always be greater than 1 to avoid the appearance of false peaks (Fig. 3D). This arbitrary choice, but based on the observed relation $P = f(N)$, represents a cut-off allowing the strict detection of real peaks. The effect of this choice is that components present in very low amounts are not detected.

Repeated smoothing-deconvolution process of overlapping chromatographic peaks

Complex chromatographic patterns were obtained by successive injections (two to five) of a radio-iodinated γ-globulin into a TSK G-3000 SW column (Fig. 4A). Different amounts were injected to obtain different complex chromatograms, leading to a signal-to-noise ratio varying from 10 to 40%. A minimum interval of 2 min was chosen between two injections. The different effects of smoothing and deconvolution on this complex chromatogram are shown in Fig. 4B and C. By using an r.r. value of greater than 1, as defined above, only real peaks corresponding to the radioactive injections were found after a smooth-deconvolution process (Fig. 4D). By using this method, false peaks were never obtained. In addition, when the signal-to-noise ratio was low, some real peak(s) could not be detected, because the effect of the r.r. was mainly to reduce the detection sensitivity, but also to increase greatly the specific detection of real peaks.

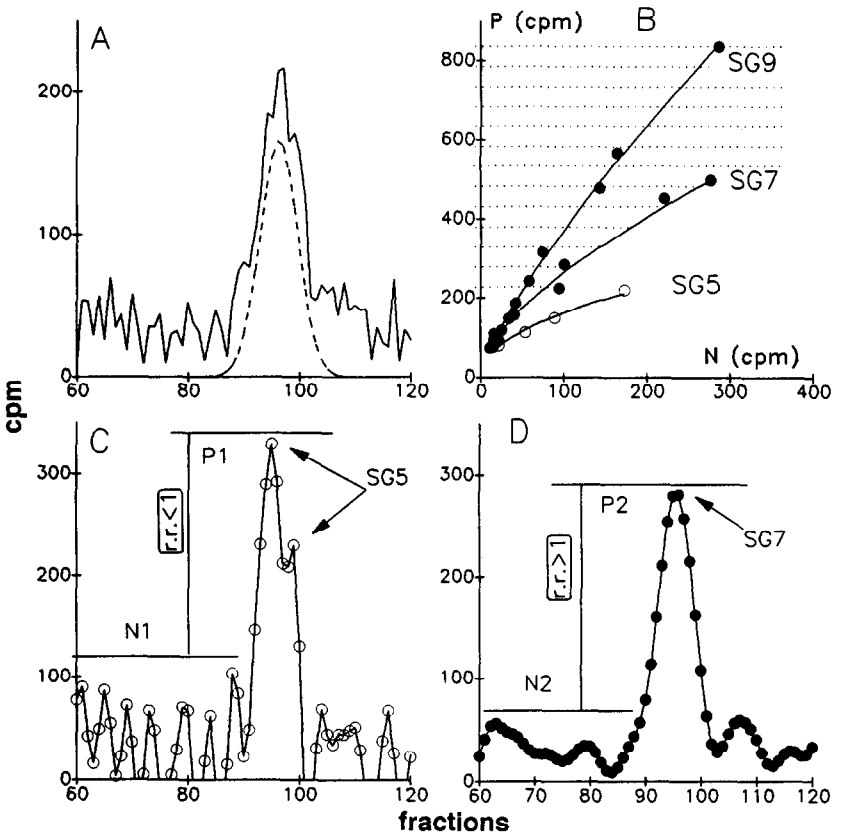

Fig. 3. Iterated smooth-deconvolution process, using three different levels of smoothing (SG_5, SG_7 or SG_9) of a Gaussian curve, in which noise has been added. (A) A 40% addition of synthetic randomly distributed noise was made with the "add noise" program to the Gaussian chromatographic curve of Fig. 2B: Gaussian curve with (———) or without (– – –) noise. The percentage of noise added was calculated relative to the Gaussian peak. (B) The curve in which noise was added was smoothed by the Savitzky–Golay function with SG_5, SG_7 or SG_9 smoothing degree. After one smooth the curve was deconvoluted with a 0.5 deconvolution filter constant. The smooth-deconvolution process was repeated several times for each degree of smoothing. After each pass of this mathematical operation, the maximum value of peak counts (P) and the maximum value of noise counts (N) were reported on the diagram. N was determined between the fractions 1 and 60, and P was determined between the fractions 61 and 180. For each smoothing degree a regression curve ($P = aN^b$) was obtained: SG_5: $P = 28.9N^{(0.493)}$; SG_7: $P = 23.2N^{(0.623)}$; SG_9: $P = 13.24N^{(0.777)}$. (C) Curve resulting from the SG_5 smoothing degree; false peaks were observed in addition with the real peak (arrows). (D) Curve resulting from the SG_7 or SG_9 smoothing degree; only one peak was observed. For a chromatogram with a maximum peak P_i and a maximum noise N_i, the ratio P_i/aN_i^b was defined as a resolution ratio (r.r.), with $a \leq 23.2$ and $b \geq 0.623$. With r.r. < 1, false peaks can be observed in addition to the real peak (C), and with r.r. > 1 only real peaks were observed (B, hatched area, and D).

Curve-fit process for overlapping chromatographic peaks

The complex chromatographic profiles obtained as described above (raw data, Fig. 4A) were further analysed by the curve-fit treatment. With this process, good results were obtained only if some parameters were first calculated. These parameters were the elution time and the peak width. The elution time was obtained from the peak position measured after a smooth deconvolution process of raw data (Fig. 4D) and the peak width was chosen as the mean peak width of all calibration proteins (3.1 ± 0.6 min). These guess values were introduced into the curve-fit program, which iterated to find the combination of peak height, position, and width that best fitted the raw data (Fig. 4E). This strategy was used to analyse twelve different complex chromatograms obtained by successive injections (two to five) of a radio-iodinated γ-globulin. A good correlation

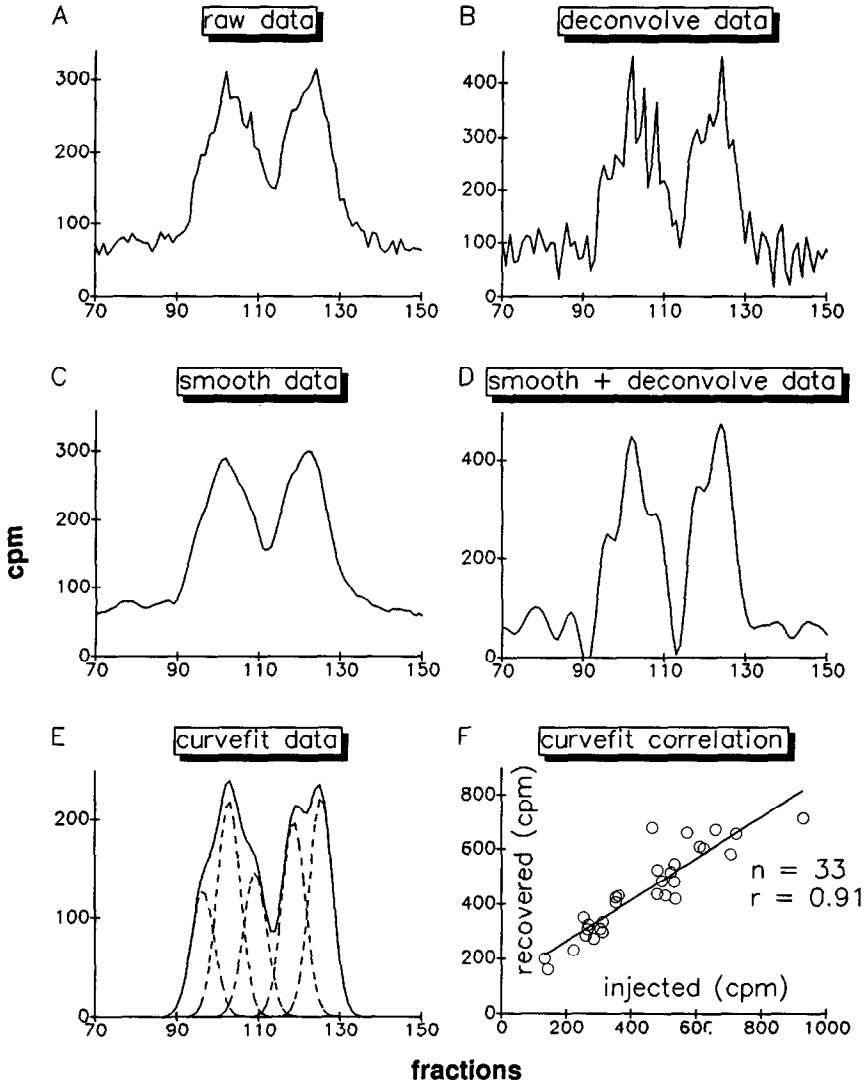

Fig. 4. Use of the smooth-deconvolution process and of the curve-fit treatment to analyse a complex chromatogram. A complex chromatogram was obtained by five successive injections of a radio-iodinated γ-globulin into a TSK G-3000 SW column. Raw counts measured by the radioactive detector (A) were deconvoluted (B), or smoothed (C) or smoothed and deconvoluted (D). Successive injections were: 0.35, 0.5, 0.25, 0.5 and 0.5 ml at times 0, 2, 4, 7 and 9 min. The smooth-deconvolution process (D), used with r.r. > 1 (see text), increased the resolution of peaks in comparison with raw data (A), and calculated elution times corresponded exactly to the real elution times of peaks. Quantitative analysis of raw data was realized by the curve-fit treatment (E). Two parameters were used with this program to calculate the peak area: the elution times calculated by the smooth-deconvolution process and the peak-width reference (3.1 ± 0.6 min) described above. Twelve complex chromatograms ($n = 33$ injections) performed as described for the raw data (A) were analysed by the curve-fit treatment. A correlation ($y = 0.76x + 109$ cpm, $r = 0.91$) was obtained between the amounts injected into the TSK G-3000 SW column and the amounts recovered by the curve-fit process (F).

(Fig. 4F) was obtained between the amounts injected (x) and recovered (y) after curve-fit analysis ($y = 0.76x + 109$ cpm; $r = 0.91$).

Smooth-deconvolution process, and curve-fit analysis of human breast cancer ER chromatograms

Fifty-six samples from human breast cancer

TABLE II

PEAK POSITION OF HUMAN BREAST CANCER ER ISOFORMS DETERMINED BY THE SMOOTH-DECONVOLUTION PROCESS ON 56 HPSEC CHROMATOGRAMS

Peak positions are expressed either as 20-s or 1-min elution fractions ± S.D. The occurrence of each individual peak in the 56 chromatograms is expressed as a percentage.

Isoform	Peak position		Peak occurrence (%)
	20-s fraction	1-min fraction	
1	75 ± 2.5	25 ± 1	94
2	83 ± 3.4	28 ± 1	97
3	90 ± 3	30 ± 1	84
4	96 ± 2.9	32 ± 1	94
5	104 ± 2.3	35 ± 1	41
6	114 ± 4	38 ± 1	41
7	119 ± 3	40 ± 1	82
8	130 ± 3	43 ± 1	25

TABLE III

MW AND R_s VALUES OF THE EIGHT ER ISOFORMS SEPARATED BY HPSEC AND ANALYSED BY THE SMOOTH-DECONVOLUTION PROCESS

For each isoform, the minima and maxima values observed are indicated.

Isoform	MW observed (kDa)			R_s observed (nm)		
	Min	Average[a]	Max	Min	Average	Max
1	600	520	470	8.1	7.6	7.3
2	390	358	320	6.6	5.9	6.2
3	260	230	200	5.2	5.4	5.6
4	195	180	160	5.1	4.9	4.8
5	140	120	110	4.4	4.2	4.1
6	72	66	61	3.5	3.4	3.3
7	50	45	41	3.1	3.0	2.9
8	26	23	20	2.8	2.7	2.6

[a] $n = 56$.

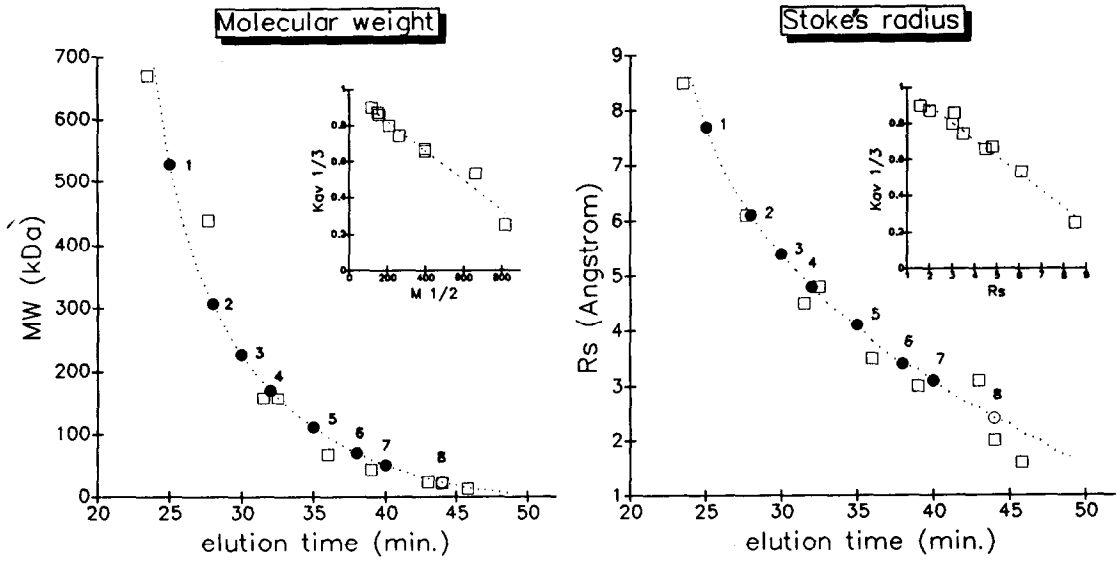

Fig. 5. Molecular masses and Stoke's radii of ER isoforms. The TSK G-3000 SW column calibration was performed using nine standard proteins: ribonuclease (MW 13.7 kDa, R_s 1.6 nm); α-chymotrypsin (MW 21.6 kDa, R_s 2.0 nm); trypsin inhibitor (MW 24 kDa, R_s 3.1 nm); ovalbumin (MW 43 kDa, R_s 3.0 nm); bovine serum albumin (MW 68 kDa, R_s 3.5 nm); aldolase (MW 158 kDa, R_s 4.8 nm); γ-globulin (MW 158 kDa, R_s 4.5 nm); ferritin (MW 440 kDa, R_s 6.1 nm); thyroglobulin (MW 669 kDa, R_s 8.5 nm). The void volume (V_0) and total liquid volume (V_t) were determined with Dextran Blue and tritiated water, respectively. The linear regression analysis of $(K_{av})^{1/3}$ *versus* R_s or *versus* $MW^{1/2}$ (insert) of the standards was plotted, where $K_{av} = (V_e - V_0)/(V_t - V_0)$, V_e being the elution volume of the protein studied. The curves "MW *versus* elution time" and "R_s *versus* elution time" (·····) were deduced, and the molecular masses and Stoke's radii of ER isoforms were estimated (Table III).

TABLE IV

RELATIVE AMOUNTS OF THE SEVEN MAJOR ER ISOFORMS DETERMINED BY THE CURVE-FIT TREATMENT OF 56 HPSEC CHROMATOGRAMS

Results are expressed as mean ± S.D.; extreme values are given in parentheses.

Isoform	Percentage of chromatogram area		
1	(8)	32 ± 9	(51)
2	(9)	19 ± 5	(30)
3	(5)	13 ± 3	(15)
4	(4)	10 ± 3	(15)
5	(3)	7 ± 2	(13)
6	(2)	7 ± 3	(17)
7	(1)	12 ± 7	(38)

cytosols labelled with [^{3}H]estradiol were injected (300–500 μl) into a TSK G-3000 SW column and eluted with PEDG-K buffer (0.4 *M* KCl). The eluted radioactivity was registered by the Flo-One radioisotope detector as a file of 180 raw counts obtained in 60 min. The presence of KCl in the elution buffer allowed the partial dissociation of ER into smaller isoforms. In order to determine the peak position of each receptor isoform, the raw data were treated by the smooth-deconvolution process as described above (with r.r. > 1), and under these conditions, eight peaks were detected, of which five were observed in more than 82% of chromatograms (Table II). A good reproducibility was observed for all peak elution times. Peak 8 was observed in only 25% of cases and represented probably a degraded form of the 70 kDa monomer. According to the peak position (elution time) of each isoform, MW values and Stoke's radii were deduced from the TSK G-3000 SW calibration curves (Fig. 5). Results, summarized in Table III, showed MWs ranging from 530 kDa (R_s = 7.7 nm) to 23 kDa (R_s = 2.6 nm). The raw data of the 56 chromatograms were further treated by the curve-fit process to measure the peak area, and thus the amount of each isoform. Calculated amounts of the seven major isoforms were expressed as percentage of the chromatogram area (Table IV). Results showed that ER isoforms dissociated into four major species (forms 1, 2, 3 and 7), corresponding to the 530, 310, 230 and 45 kDa isoforms. Furthermore, the dissociation process varied greatly from tumour to tumour, with ER being fully or quite not dissociated after chromatography (Fig. 6). The presence and amount of the 45 kDa isoform was often related to high ER dissociated forms. Varying concentrations of KCl (0–1 *M*) were used with a low ER dissociated form in order to study the receptor dissociation process by the computer program. After incubation for 1 h with KCl, and chromatographic analysis, the ER polymeric form (530 kDa; 7.7 nm) dissociated into six different isoforms (forms 2–7), as detected by the smooth-deconvolution analysis (Fig. 7).

Similarly, the smoothing-deconvolution procedures were used for the study of breast cancer progesterone receptor isoforms from 45 cytosols. PR was more resistant to KCl dissociation than ER, and a pretreatment of the receptor by KCl (0.4 *M*; 5 h) was used to obtain mild dissociation of PR isoforms. These pretreated PR were injected in the column and chromatograms were processed as previously for ER. Under these conditions, eight PR isoforms were obtained, with MW ranging from 680 kDa (R_s = 8.6 nm) for the polymeric structure, to 50 kDa (R_s = 3.1 nm) for the most dissociated form (Table V). As for ER, wide differences between tumours were observed in the dissociation kinetics of PR.

DISCUSSION

Chromatographic analysis of breast cancer ER polymeric forms, by HPSEC in the presence of high salt concentrations, yields different complex chromatograms that are tumour-specific, ER being either resistant or not to the dissociation effect of the salt. These chromatograms represent the elution of several ER isoforms obtained by the dissociation effect of KCl, and overlapping of peaks can hide some of eluted isoforms.

Different computer programs exist to improve the detection of peaks in complex chromatograms, and the program we have used is powerful

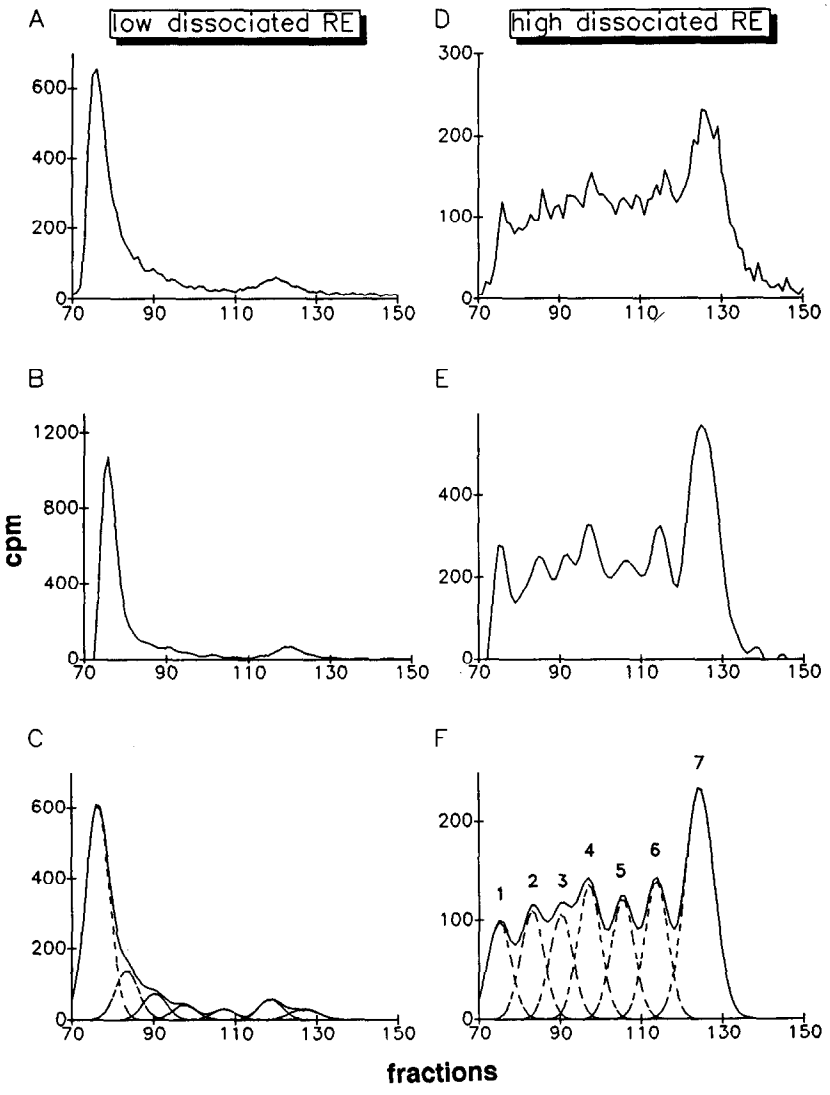

Fig. 6. HPSEC analysis of the dissociation effect of 0.4 *M* KCl on the polymeric form of human breast cancer estrogen receptors. Cytosols ($n = 56$) were incubated with [^{3}H]estradiol and treated with Dextran-coated charcoal, as described in Experimental. Aliquots were chromatographed on a TSK G-3000 SW column, and the elution of receptor isoforms was performed with PEDG-K buffer (containing 0.4 *M* KCl). The radioactivity of the eluate was recorded with a radioisotope detector (Flo-One) fitted on line to the column outlet. Raw data from weakly dissociated (A) and highly dissociated (D) ER forms were treated by a smooth-deconvolution strategy (B and E) for the determination of the ER isoform elution times. Raw data were then treated by the curve-fit program, for quantitative analysis of ER isoforms (C) and (F).

enough to analyse chromatographic peaks and enhance resolution by mathematical functions involving smoothing and deconvolution. Elution studies of repeated injections of proteins through the TSK G-3000 SW column, and analysis of the complex constructed chromatograms by the computer program, were performed first to define a general strategy to be used with unknown chromatographic patterns, particularly to prevent detection of false peaks. By using standard proteins, it was observed that the peak width of eluted proteins with MWs ranging from 669 kDa (thyro-

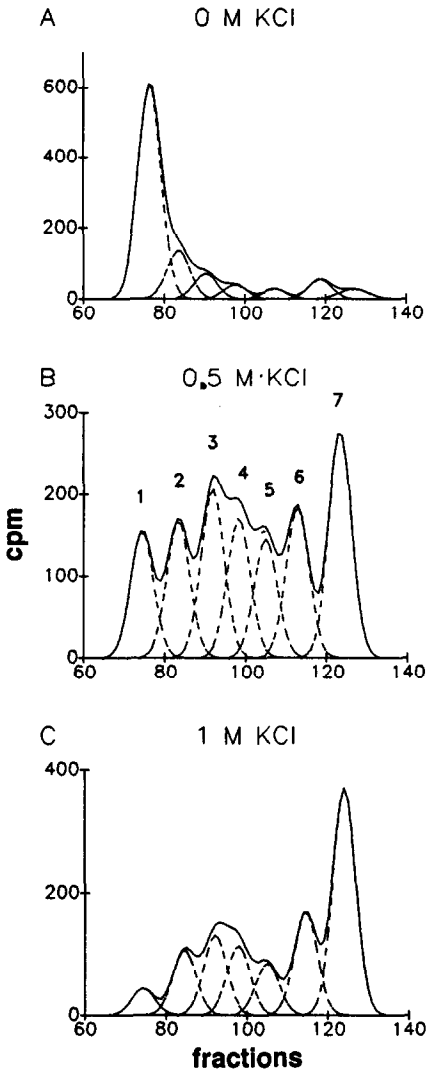

Fig. 7. HPSEC analysis of the KCl-induced dissociation of breast cancer ER isoforms, and curve-fit process of raw data. Cytosols were incubated with [^{3}H]estradiol and treated with Dextran-coated charcoal, as described in Experimental. Aliquots were treated with various amounts of KCl, ranging from 0 to 1 *M*, for 1 h in an ice-bath, and then immediately chromatographed on a TSK G-3000 SW column filled with PEDG-K (0.4 *M* KCl) buffer. The radioactivity of the eluate was recorded with a radioisotope detector (Flo-One) fitted on line to the column outlet. Raw data were treated by a smooth-deconvolution strategy for the determination of the ER isoform elution volumes. Raw data were then treated by the curve-fit treatment for quantitative analysis of ER isoforms (– – –).

globulin) to 13.7 kDa (ribonuclease) was relatively constant (3.1 ± 0.6 min), and this value was used as reference for the curve-fit treatment of the

TABLE V

MW AND R_s VALUES OF THE EIGHT PR ISOFORMS SEPARATED BY HPSEC AND ANALYSED BY THE SMOOTH-DECONVOLUTION PROCESS

For each isoform, the minima and maxima values observed are indicated.

Isoform	MW observed (kDa)			R_s observed (nm)		
	Min	Average[a]	Max	Min	Average	Max
1	740	680	620	9.0	8.6	8.1
2	660	600	550	8.6	8.1	7.6
3	470	420	380	7.2	6.9	6.6
4	268	246	227	5.6	5.4	5.2
5	194	178	165	5.1	4.9	4.7
6	128	119	110	4.3	4.2	4.1
7	94	86	80	3.8	3.7	3.6
8	55	50	45	3.2	3.1	3.0

[a] $n = 45$.

ER chromatograms. Complex chromatograms were easily constructed by injecting different known amounts of a protein at different time intervals: the program was tested in order to find again the elution time and the amount of the different "components" injected into the column. In particular, two functions were successfully used: the SG smoothing function [9] and the FSD function [12]. The SG function both smooths a peak and increases its width, whereas the FSD function acts inversely. Because these two functions use different mathematical processes, their opposite effects create interesting transformations of the chromatographic raw data. In particular, when overlapping peaks were first identified as a single large peak, the smooth-deconvolution process was able to separate the real peaks out. Thus, treatment of chromatographic raw data by this smooth-deconvolution operation is a powerful method for the improvement of the resolution of overlapping peaks. To avoid the possibility of aberrant results (false peak detection), a strategy was developed, including the use of SG and FSD functions, in addition to a calculated r.r. With r.r. values higher than 1, all false peaks due to noise were avoided. The curve-fit program allows the

measurement of peak areas, and, by using standard proteins, 80% of the amount injected was regularly recovered. This suggests that the Gaussian curve obtained after the curve-fit treatment reflects well the original eluted profile. Furthermore, the use of this program with complex chromatograms showed a good correlation between the amounts of component injected and recovered after HPSEC.

Human breast cancer ER is found in cytosolic extracts as a heteropolymer: its quaternary structure is not fully understood. Different proteins (hsp90, hsp70, p59) are associated with ER [13], and other proteins (MW 54, 50, and 23 kDa) have been co-isolated with PR [14], but the association of these proteins with the ER-hsp90-hsp70-p59 complex has not been convincingly demonstrated.

The salt effect of KCl dissociates the ER polymeric structure into different isoforms, from highly polymeric to highly dissociated forms, allowing the detection of different isoforms. In these conditions, breast tumour ER chromatograms were different from tumour to tumour, with ER being fully or quite not dissociated after chromatography. This strategy has allowed the detection of eight ER isoforms, with MW and R_s ranging from 530 to 23 kDa, and from 7.7 to 2.6 nm, respectively; these isoforms were recovered in variable amounts for each tumour. Isoform 8 is recovered in only 25% of cases, and is probably a proteolytic fragment of ER. Isoform 7 (45 kDa) may correspond to the ER monomeric form detected recently by other laboratories [15, 16]: a 52 kDa isoform, the mRNA of which has lost exon 7 of the 70 kDa ER mRNA, has been described [15], and a 47 kDa isoform, which binds the anti-estrogen tamoxifen-aziridine, was found in variable amounts in different tumours, in addition to the monomeric 70 kDa ER isoform [16]. The presence and amount of the 45 kDa isoform is related to highly dissociated forms of ER, suggesting that the quaternary structure of the polymeric 45 kDa ER is less stable than that of the 70 kDa ER. If this is the case *in vivo*, the presence in a given tumour of a labile ER, containing a truncated and altered ER monomer, could explain why some ER+ tumours do not respond to endocrine therapy.

Study of PR isoforms is complicated by the fact that two different polymeric forms co-exist on the tumour, one containing a 120 kDa monomer, and the other a 90 kDa monomer, both associated with heat-shock proteins, like ER. So, during dissociation, isoforms originating from either polymer migrate very close to each other, so that some of the "isoforms" observed (Table V) probably represent two different species. This illustrate clearly the fact that the column and the computer program have their own limitations.

For further studies, HPSEC is an excellent tool for isoform analysis of breast cancer ER and PR, allowing a better definition of receptor structure and functionality, and thus giving new insights for clinical therapy.

ACKNOWLEDGEMENT

This work was supported by research grants from the Comité du Calvados de la Ligue Nationale Française contre le Cancer.

REFERENCES

1 W. L. McGuire, *Semin. Oncol.*, 5 (1978) 428.
2 J. L. Wittliff, *Cancer*, 46 (1980) 2953.
3 M. R. Sherman and J. Stevens, *Annu. Rev. Physiol.*, 46 (1984) 83.
4 M. R. Sherman, L. A. Pickering, F. M. Rollwagen and L. K. Miller, *Fed. Proc., Fed. Am. Soc. Exp. Biol.*, 37 (1978) 167.
5 J. Goussard, C. Lechevrel, P.-M. Martin and G. Roussel, *Cancer Res.*, 46 (1986) 4282s.
6 R. Hähnel and E. Twaddle, *J. Steroid Biochem.*, 10 (1979) 95.
7 O. H. Lowry, N. J. Rosebrough, A. L. Farr and R. J. Randall, *J. Biol. Chem.*, 193 (1951) 265.
8 H. Cren, C. Lechevrel, G. Roussel and J. Goussard, *J. Steroid Biochem. Molec. Biol.*, 39 (1991) 519.
9 A. Savitzky and M. J. E. Golay, *Anal. Chem.*, 36 (1964) 1627.
10 J. Steiner, Y. Termonia and J. Deltour, *Anal. Chem.*, 44 (1972) 1906.
11 H. Madden, *Anal. Chem.*, 50 (1978) 1383.
12 P. R. Griffiths and G. L. Pariente, *Trends Anal. Chem.*, 5 (1986) 209.
13 S. Y. Tsai, M.-J. Tsai and B. W. O'Malley, in M. G. Parker (Editor), *Nuclear Hormone Receptors*, Academic Press, London, 1991, p. 103.
14 D. F. Smith, L. E. Faber and D. O. Toft, *J. Biol. Chem.*, 265 (1990) 3996.
15 S. A. W. Fuqua, S. D. Fitzgerald, D. C. Allred, R. M. Elledge, Z. Nawaz, D. P. McDonnell, B. W. O'Malley, G. L. Greene and W. L. McGuire, *Cancer Res.*, 52 (1992) 483.
16 S. Jozan, A.-M. Julia, A. Carretie, N. Eche, V. Maisongrosse, B. Fouet, B. Marques and J. F. David, *Breast Cancer Res. Treat.*, 19 (1991) 103.