

Refinement of the C222₁ Crystal Form of Oxidized Uteroglobin at 1.34 Å Resolution

I. Morize¹, E. Surcouf¹, M. C. Vaney¹, Y. Epelboin¹, M. Buehner²
F. Fridlansky³, E. Milgrom³ and J. P. Mornon^{1†}

¹ *Laboratoire de Minéralogie-Cristallographie
associé au CNRS, Universités Paris 6 et Paris 7
Tour 16, 4 Place Jussieu, 75230 Paris Cedex 05, France*

² *A.G. Roentgenstrukturanalyse, Universitaet Wuerzburg
Am Hubland, D-8700 Wuerzburg, Federal Republic of Germany*

³ *Hormones et Reproduction
(INSERM U135), Faculté de Médecine Paris Sud.
78 Avenue du Général Leclerc, 94275 Kremlin Bicêtre
Cedex, France*

(Received 13 May 1986, and in revised form 19 December 1986)

The structure of uteroglobin, a progesterone binding protein from rabbit uterine fluid, was determined and refined at 1.34 Å resolution to a conventional *R*-factor of 0.229. The accuracy of the co-ordinates is estimated to be 0.15 Å. The isotropic temperature factor of individual atoms was refined and its average value is 11.9 Å² for the 548 non-hydrogen atoms of the protein monomer. A total of 83 water molecules was located in difference electron density maps and refined, first using a constant occupancy factor of 1 and then variable occupancy, the final $\langle Q \rangle$ being 0.63. The mean temperature factor of the water oxygen atoms is 26.4 Å². Uteroglobin is a dimer and its secondary structure consists of four α -helices per monomer that align in an anti-parallel fashion. There is one β -turn between helix 2 and helix 3 (Lys26 to Glu29); 76% of the residues are part of the α -helices. In the core of the dimeric protein molecule, between the two monomers that are held together by two disulfide bridges, we have observed a closed cavity. Its length is 15.6 Å and its width is 9 Å; 14 water molecules could be positioned inside. In the "bottom" part of the protein, near the C terminus, we have observed a smaller cavity, occupied by two water molecules. The calculation of the molecular surface revealed four surface pockets whose possible functional implications are discussed below.

1. Introduction

Uteroglobin is a protein in the uterine secretion of the rabbit (Beier, 1968; Bullock *et al.*, 1976; Beato *et al.*, 1980; Noske & Feigleson, 1976; Bullock, 1977). Its synthesis is induced by progesterone and progestins (Atger *et al.*, 1980; Savouret & Milgrom, 1983; Lopez de Haro & Nieto, 1983). Low concentrations of uteroglobin are also found in the lungs where it may be controlled, at least in some situations, by glucocorticoids (Lopez de Haro & Nieto, 1983, 1985).

The function of uteroglobin is not clearly understood, but it is known that it binds progesterone and metabolites of progesterone. Since uteroglobin has been purified and crystallized, it

offers an interesting model with which to study the three-dimensional structure of steroid-binding proteins.

The molecular weight of uteroglobin is 15,782 and it is made of two identical polypeptide chains of 70 amino acids each (cf. Fig. 1). Two disulfide bridges (Cys3–Cys69' and Cys69–Cys3') connect these two chains in the oxidized *in vitro* form. *In situ*, the cysteine residues are reduced. Binding and exchange of ligands takes place in the reduced state only (Tancredi *et al.*, 1982). Attempts to crystallize the reduced molecule, or to crystallize a uteroglobin–ligand complex, have been unsuccessful.

The main characteristics of its amino acid sequence (Ponstingl *et al.*, 1978; Popp *et al.*, 1978; Atger *et al.*, 1979) is the lack of tryptophan and the relatively small number of aromatic residues: 1 tyrosine and 2 phenylalanine only (cf. Table 1).

Our first structural study allowed us to define the

† Author to whom correspondence should be addressed.

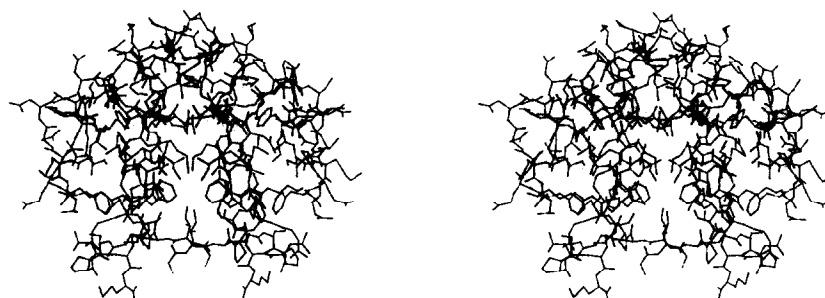


Figure 1. Stereo view of uteroglobin along the *c*-axis.

structure of uteroglobin at 2.2 Å resolution (Mornon *et al.*, 1980). We observed the existence inside the protein of a central cavity not accessible from the solvent. Its size, shape and chemical characteristics suggest that it might be the steroid binding pocket. Moreover, the presence of a large volume of electron density in the isomorphous replacement map led us to wonder whether the protein has bound progesterone or a similar steroid molecule in this cavity.

The present study was aimed at refining the *C*222₁ form of uteroglobin, to establish its three-dimensional structure at higher resolution, to study further characteristics of the internal hydrophobic cavity and also to analyze the protein for the existence of other possible binding sites. Finally, since six crystalline forms of uteroglobin have been observed (Mornon *et al.*, 1978, 1979; Buehner & Beato, 1978; Buehner *et al.*, 1982) it was appealing to try to understand better the flexibility and mobility of uteroglobin using its crystalline polymorphism. The crystal structure of the *P*2₁ form (Mornon *et al.*, 1979) which contains one dimer per asymmetric unit and was solved by isomorphous replacement (Mornon & Bally, unpublished results), is now at the end of the refinement using 1.7 Å data (study performed by R. Bally and J. Delettré).

In this paper we describe fully the structural analysis of uteroglobin as it is present in the *C*222₁ form.

2. Experimental Procedures

(a) Crystallization and data collection

For crystallization the vapor diffusion method was used, at room temperature, with ammonium sulfate as the precipitating agent. Yellow crystals of good size, about 0.5 mm × 0.6 mm × 0.3 mm, were obtained in space group *C*222₁ with unit cell dimensions *a* = 50.86 Å, *b* = 52.22 Å, *c* = 47.28 Å and one monomer per asymmetric unit (Mornon *et al.*, 1978). The yellow color of the crystals cannot be explained. It is most probably a packing artefact, because different crystal forms exhibit different shades of yellow, as well as white.

The present refinement was based on data collected from 4 crystals of the native protein, while the preliminary structure analysis was carried out using data sets of native crystal I and those of heavy-atom derivatives UO₂, PtCl₄(I) and PtCl₄(II).

Intensities were collected using a Philips PW1100 automatic diffractometer in our laboratory (routine Stepscan). The tube has been operated at 50 kV and 20 mA, using monochromatized CuKα radiation.

X-radiation damage was monitored by periodic measurement of several reflections distributed in all θ

Table 1
Sequence and composition of one monomer of uteroglobin

1	5	10	15	20	25	30
G-I- C- P- R- F- A- H- V- I- E- N- L- L- L- G- T- P- S- S- Y- E- T- S- L- K- E- F- E- P-						
	35	40	45	50	55	60
D- D- T- M- K- D- A- G- M- Q- M- K- K- V- L- D- S- L- P- Q- T- T- R- E- N- I- M- K- L- T-						
	65	70				
E- K- I- V- K- S- P- L- C- M						

Type	No.	Residue	Type	No.	Residue
A (Ala)	2	7 37	M (Met)	5	34 39 41 57 70
R (Arg)	2	5 53	F (Phe)	2	6 28
N (Asn)	2	12 55	P (Pro)	5	4 18 30 49 67
D (Asp)	4	31 32 36 46	S (Ser)	5	19 20 24 47 66
C (Cys)	2	3 69	T (Thr)	6	17 23 33 51 52 60
E (Glu)	5	11 22 27 29 54	W (Trp)	0	
Q (Gln)	3	40 50 61	Y (Tyr)	1	21
G (Gly)	3	1 16 38	V (Val)	3	9 44 64
H (His)	1	8			
I (Ile)	4	2 10 56 63			
L (Leu)	8	13 14 15 25 45 48 59 68			
K (Lys)	7	26 35 42 43 58 62 65			

Table 2
Measurement of native crystals C222₁

Crystal number:	1	2	3	4
<i>a</i> (Å)	50.86	50.93	51.01	50.90
<i>b</i> (Å)	52.22	52.47	52.31	52.34
<i>c</i> (Å)	47.28	47.28	47.17	47.16
Resolution (Å)	22.0–1.64	2.06–1.54	1.82–1.45	1.73–1.34
Number of reflections	7984	6675	5853	7503
Max. decay correction on intensities (%)	11	17	22	38
<i>R</i> _m	Ref.	0.107	0.129	0.089

*R*_m, mean deviation of data from crystals 2, 3 and 4 with respect to crystal 1; $R_m = (\sum (I - I_1)) / \sum I_1$.

ranges and/or by remeasurement of the first reflections at the end of the data collection of each crystal.

The observed structure factors ($|F_o|$) were corrected for Lorentz factor, polarization and radiation damage. In addition, absorption corrections were applied to crystal 1.

The maximum values of the decay correction are shown in Table 2; they are a function of time only, and we also note the expected behavior in $\sin \theta$: greater damage corresponds to higher resolution.

The intensity data collected from different crystals were reduced to a common scale and merged together into a single set of data.

During this step, we have noticed small differences of the largest structure factors and relatively large differences of the smallest ones, between the different crystals. To take into account this distribution, the reflections for which the agreement between the different crystals was good were weighted higher in the refinement than those for which the deviation was large.

(b) The 2.2 Å resolution starting model†

The electron density based on multiple isomorphous replacement was initially interpreted by reading approximate co-ordinates directly from mini maps, using the published model (Mornon *et al.*, 1980).

After correction of 2 side-chains (the assignment of one piece of density was wrong as revealed by comparison with the structure in crystal form *P2₁2₁2*) the rough co-ordinates were regularized by the procedure of Fitzwater & Scheraga (1980). This method was employed in Wuerzburg where neither a good graphics system nor FRODO is available.

After some difficulties of initiation, the procedure worked smoothly; regularization was done in a zone that moved along the polypeptide chain. In the first run through the molecule, a zone width of about 5 amino acids was mostly used, and the zone was in general shifted by 1 amino acid per main cycle (each main cycle consisted of 7 least-squares cycles), so that every amino acid was refined in at least 4 main cycles.

After completion of the first run, which took 90 main cycles, the root-mean-square (r.m.s.) deviation of the model co-ordinates from the rough co-ordinates was 0.50 Å. At this point the sulfur–sulfur distance of the

disulfide bridges was manually adjusted (from 2.5 Å to 2.0 Å), and regularization was completed in a second run through the molecule using larger zone-widths and shifts. The final r.m.s. deviation of co-ordinates was 0.48 Å, and the sulfur–sulfur distance had gone up to 2.3 Å again.

The regularized model was then corrected by reinterpreting the multiple isomorphous replacement electron density using program FRODO (Jones, 1982). Amino acids were fitted to the electron density, and the stereochemistry was then restored using the procedure of Hermans & McQueen (1974), which is a part of FRODO.

Model building was completed by a final run of the Hermans & McQueen (1974) procedure using a zone 6 amino acids wide, shifted by 3 amino acids per procedure call.

The model was further improved by joint X-ray and energy refinement (EREF) (Jack & Levitt, 1978).

The FRODO model of uteroglobin was subjected to 9 cycles of EREF at first, then Cys3–Cys69 and some other residues were rebuilt and 4 more cycles of EREF were done. The first EREF cycle after each model building was executed without X-ray contribution. Only positions were refined at that stage, and no weights were used. The overall statistics of refinement of the model are given in Table 3 and, in Fig. 2, the dependence of residual and correlation coefficient on the resolution is shown. It is quite obvious that, for the beginning stages, the correlation coefficient is more discriminating than the residual.

3. Refinement

(a) Restrained parameter least-squares refinement

The restrained parameter least-squares procedure of Hendrickson & Konnert (1980a) was used to refine the structure. Furey's version of the program using an array processor (Furey *et al.*, 1982) was implemented on a DPS6 CII-HB computer (16 bits) and a Floating Point Systems AP100 array processor. Some technical modifications had to be made due to limitations of the computer.

The execution time per refinement cycle was approximately 25 minutes for about 14,000 reflections, 548 non-hydrogen atoms and 83 water molecules (2608 parameters).

The different restraints employed were: covalent bond lengths, distances corresponding to interbond angles, planarity of peptide bonds and aromatic rings, chirality, van der Waals' contacts, hydrogen

† This work was performed at the University of Wuerzburg. For FRODO model building and EREF refinement the facilities of the Max-Planck-Institut fuer Biochemie at Martinsried (Vector General 3400 display and VAX-11/782) were kindly made available to us by Professor Robert Huber.

Table 3
Course of pre-refinement of the 2.2 Å uteroglobin model

Model	R (%)	Correlation	Remarks
XR	50.31	0.3653	Rough X-ray co-ordinates
FS	49.10	0.4171	Fitzwater & Scheraga-regularized
VG	43.54	0.5747	FRODO-built on Vector General display
E5	37.05	0.6861	Energy + X-ray refinement, cycle 5
E8	35.43	0.7083	Energy + X-ray refinement, cycle 8
E12	33.23	0.7402	Energy + X-ray refinement, cycle 12

$$R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \times 100.$$

$$\text{Correlation} = \frac{n \sum (|F_o| \times |F_c|) - \sum |F_o| \times \sum |F_c|}{\sqrt{((n \sum |F_o|^2) - (\sum |F_o|)^2) \times (n \sum |F_c|^2) - (\sum |F_c|)^2}}.$$

Structure factors were calculated by FFT (fast Fourier transform) from density expanded co-ordinates ($B = 25 \text{ Å}^2$) and were fitted to the observed data by a 2-parameter least-squares procedure minimizing $\sum (|F_o| - |F_c|)^2$.

bond distances, conformational torsion angles and isotropic thermal factors. All of them were used with weighting factors as suggested by Hendrickson & Konnert (1980b).

(b) Refinement survey

The refinement process was divided into four stages. A total of 94 cycles was performed.

Table 4 gives a summary of the refinement, and Figure 3 shows the variation in the R -factor.

As the starting model was already pre-refined, it was not necessary to idealize it. We used a structure factor weight $\sigma(f) = \frac{1}{2} \langle |F_o| - |F_c| \rangle$ right from the beginning. This allowed us to maintain a balance between restraint weights and structure factor weights.

In the first stage (18 cycles), data were gradually

added up to the 1.34 Å resolution limit and no manual revision of the model was done.

At this point of refinement an R -factor of 0.264 was reached for data from 5 to 1.34 Å resolution.

During the second stage, we examined "fragmented" difference maps, i.e. $(|F_o| - |F'_c|) \exp(\varphi_c)$ syntheses where $|F'_c|$ is the calculated structure factor of a partial model from which a fragment of about ten residues has been deleted. For the systematic inspection of the ΔF maps, we first used a Benson plotter and later a graphics display system PS300 from Evans & Sutherland (with version 0 of the program MANOSK).

The most important modifications of the model concerned the N and C termini (from Gly1 to Pro4 and from Pro67 to Met70), where the two disulfide bridges are located. These parts had been unstable in the starting model as indicated by the persistent distortion of the S-S distance.

Other corrections were necessary for different

Table 4
Course of uteroglobin refinement

Stage no.	I		II		III		IV	
	1	18	32	47	64	78	88	94
Cycle no.								
Bond distance (Å)	0.017	0.024	0.019	0.016	0.018	0.024	0.024	0.013
No. bond dist. $> 2\sigma(d)^\dagger$	87	179	158	116	115	169	183	45
Angle distance (Å)	0.045	0.065	0.066	0.060	0.057	0.060	0.059	0.041
Chirality (Å ³)	0.271	0.314	0.300	0.291	0.247	0.291	0.295	0.204
Planarity (Å)	0.048	0.017	0.014	0.019	0.017	0.019	0.014	0.008
Torsion 1 (°)‡	10.3	9.6	11.9	10.9	4.0	3.3	4.7	5.1
Torsion 2 (°)‡	21.5	21.5	21.9	18.3	18.6	19.3	19.6	19.3
$\langle F_o - F_c \rangle$	13.41	6.52	6.65	6.38	4.50	3.82	4.05	3.60
$\langle \sigma(F_c) \rangle$	6.16	3.55	3.83	7.61	3.74	2.41	3.23	4.63
$\langle \text{Co-ord. shift} \rangle$ (Å)	0.161	0.346	0.283	0.139	0.234	0.142	0.097	0.086
$\langle B \rangle$ protein (Å ²)	11.9	10.73	8.86	8.16	10.14	13.07	13.22	11.91
Resolution range (Å)	5.0-2.0	5.0-1.34	5.0-1.34	5.0-1.34	5.0-1.34	4.6-1.47	all	all
No. reflections	3502	7992	7992	7992	12,388	10,197	14,361	14,227
No. parameters	2189	2194	2141	2285	2433	2525	2628	2608
No. protein atoms	547	548	535	531	548	548	548	548
No. solvent atoms	0	0	0	40	60	83	87	83
R	0.371	0.264	0.269	0.258	0.233	0.199	0.241	0.229

$^\dagger \sigma(d) = 0.020 \text{ Å}$ for covalent bonds.

‡ Torsion 1, for peptide bond Ca-C-N-Ca. Torsion 2, for asymmetric carbon.

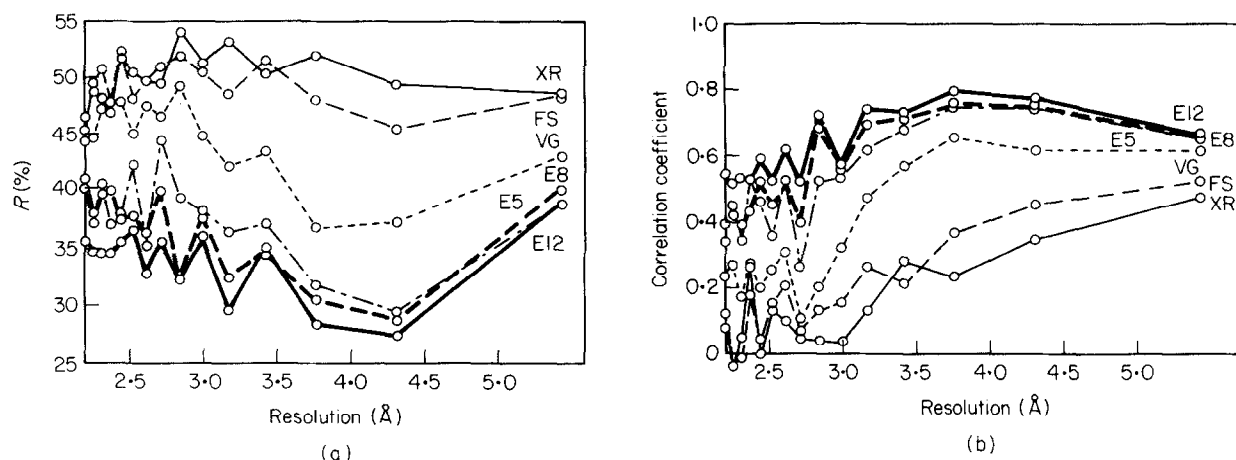


Figure 2. Course of the pre-refinement of uteroglobin. (a) *R*-factors and (b) correlation coefficients of the different stages are plotted as a function of the resolution.

side-chains (His8, Glu11, Glu22, Leu25, Lys26, Glu27, Lys35, Gln50, Arg53, Glu61 and Lys65) but the fit of the backbone was in general good enough to keep shifts to a minimum. At the same time, 40 water molecules could be located and were included in the refinement while 17 protein atoms were taken out of the model.

At the end of this stage the *R*-factor was 0.258 for 7992 reflections from 5 to 1.34 Å resolution, where $|F_o| > 2\sigma(|F_o|)$.

To complete our model and improve the solvent structure, we included in the refinement those structure factors for which $|F_o| > \sigma(|F_o|)$.

The first consequence was a large increase in the *R*-factor, followed by a steep descent during the following cycles.

Two more ΔF maps were calculated and analyzed. All protein atoms previously missing were found and 43 water molecules could be added.

The *R*-factor reached 0.219 for 12,388 reflections

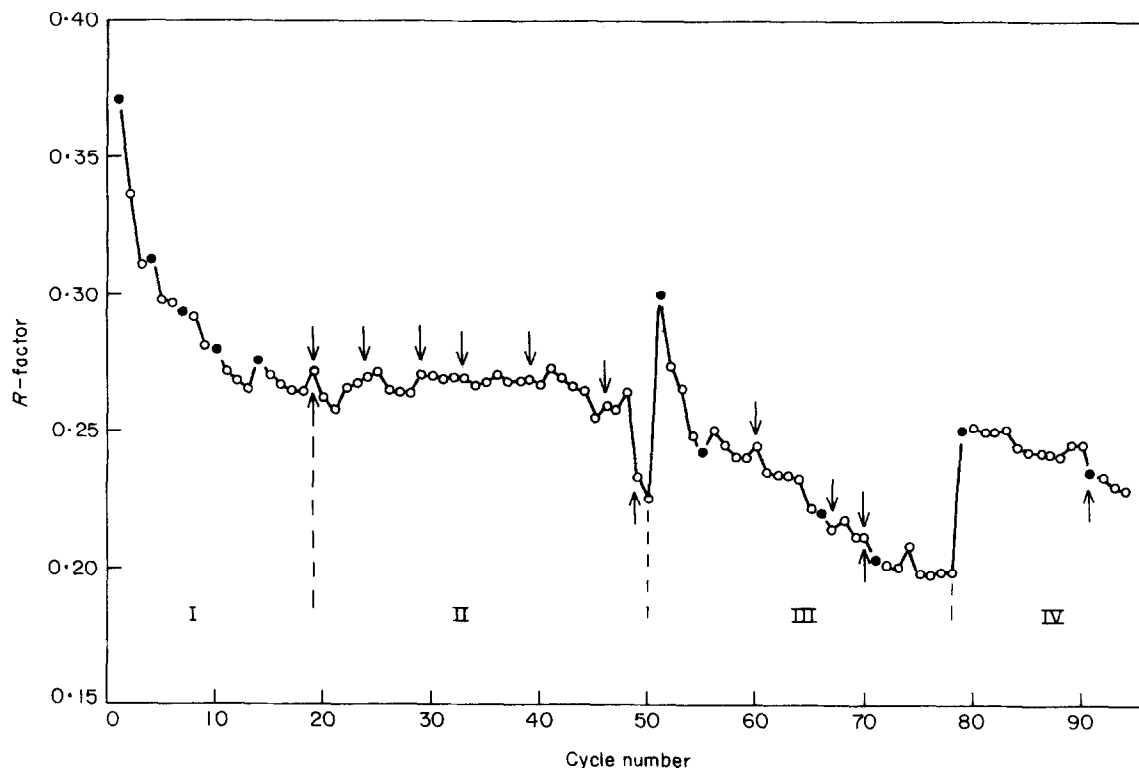


Figure 3. Progress of the refinement. The variation of the *R*-factor is plotted as a function of the cycle number. Stage I, gradual addition of the observed structure factors up to the 1.34 Å resolution limit. Stage II, inspection of "fragmented" difference ($|F_o| - |F'_c|$, φ_c) map ΔF map). $|F'_c|$ is the calculated structure factor of the partial model (with about 10 residues excluded). Stage III, addition or removal of the weakest reflections; 3 more maps were studied to complete the solvent structure. Stage IV, introduction of all reflections, of the variable occupancy factor for the water molecules and also introduction of theoretical hydrogen atoms for the protein. $\uparrow$: New scheme of weight on structure factors; $\downarrow$: calculation of ΔF map; (●) variation in the number of reflections.



Figure 4. Electron density map ($|F_o|, \phi_\gamma$) calculated around Tyr21, at the end of the refinement.

from 4.6 to 1.5 Å resolution, where $|F_o| > \sigma|F_o|$ and 0.199 for 10,197 reflections with $|F_o| > 3\sigma(|F_o|)$ in the same resolution range.

Up to this stage, the occupancy of the water molecules had not been allowed to vary. Then, in the fourth stage, we included this last parameter as a variable. During this stage, our weighting strategy was aimed at improving the stereochemistry of the model. For this reason, we included 556 theoretical hydrogen atoms as a fixed part of the structure factors

$$(|F_c| = \text{sqrt}((A_c + \text{Fixa})^{**2} + (B_c + \text{Fixb})^{**2})).^\dagger$$

In this way, the individual thermal factors B were corrected or, more exactly, could take on meaningful values.

The discontinuity in the R -factor plot between stages 3 and 4 is the result of the addition of all the weak structure factors. At the end of the refinement we scrutinized large differences between $|F_o|$ and $|F_c|$ and excluded 147 reflections for which the discrepancy between $|F_o|$ and $|F_c|$ could be traced to measurement errors. This improved the R -factor by 1%.

The final R -factor value was 0.229 for all reflections (14,227) and 0.21 for the most significant ones (8237).

The total mean co-ordinate shift was 0.59 Å, with 7.18 Å as the largest displacement for NZ of Lys43 and 0.022 Å as the smallest for Ca Ile10.

4. Refinement Results and Discussion

(a) Accuracy

Superposition of the R -factor curve with theoretical curves for different mean positional

[†] A_c and B_c are the contributions of C, N, O, S atoms, allowed to vary; Fixa and Fixb those of fixed H atoms.

Table 5
Bond geometry in uteroglobin

Type	No.	Mean value	r.m.s. dev.	Min.	Max.
N-Ca	70	1.472 Å	0.013	1.447	1.512
Ca-C	70	1.528 Å	0.009	1.508	1.548
(=O)	70	1.245 Å	0.014	1.211	1.281
C-N	69	1.326 Å	0.012	1.306	1.359
Ca-Cb	67	1.533 Å	0.021	1.494	1.597
N-Ca-C	70	110.62°	2.662	104.43	117.65
Ca-C=O	70	119.73°	2.997	110.56	125.88
Ca-C-N	69	115.63°	2.000	112.12	121.34
O=C-N	69	124.53°	2.446	119.06	131.39
Ca-C-N-Ca	69	176.12°	3.484	164.63	180.00

errors indicates a mean co-ordinate error of 0.15 Å (Luzatti, 1952).

The actual accuracy is certainly better for the rigid parts of the molecule and worse for the flexible parts, especially for some side-chains exposed to the solvent and/or with high temperature factors.

The variation in the R -factor as a function of $\sin(\theta)/\lambda$ and as a function of $|F_o|$ is normal. We observe that the lowest values are in the 4.6 Å to 2.0 Å resolution range, where we also find the highest values of the average structure factor amplitudes. Beyond 2.0 Å resolution we have the weakest reflections, i.e. reflections affected most by errors of measurement, then the R -factor increases. Inside 4.6 Å resolution the increase in the R -factor is the consequence of the omission of solvent contributions.

The r.m.s. deviation of model parameters from their ideal values are listed in Table 4 (cycle 94). It can be seen in particular that the r.m.s. deviation of bond length from ideality is 0.013 Å and only 45 of the 556 bonds deviate by more than 0.040 Å ($=2\sigma(d)$, where $\sigma(d)$ is the standard deviation).

Table 5 shows the deviation of bond distances and bond angles of the peptide main chain from their restrained values. The mean value of each parameter is in good agreement with the value of the restraint.

All data mentioned above are important for the assessment of the stereochemistry of the protein and the result of the refinement, but another way to estimate the quality of the model is to examine Fourier maps and difference Fourier maps.

The remaining peaks in the difference Fourier map are associated with second sites of disordered atoms (e.g. OH Ser66, water molecules).

Figure 4 shows the electron density map ($|F_o|, \phi_c$) in the region of Tyr21, which is situated near the central cavity and whose thermal factor is 7.6 Å². We can appreciate the high resolution (1.34 Å) of the aromatic ring of Tyr21, for example, where the carbon atoms separated by 1.40 Å are quite well defined, while the oxygen atom and carbon atom separated by 1.20 Å, in the back of Tyr21, are not so well resolved. Carbon atoms separated by 1.54 Å are often very well resolved throughout the map.

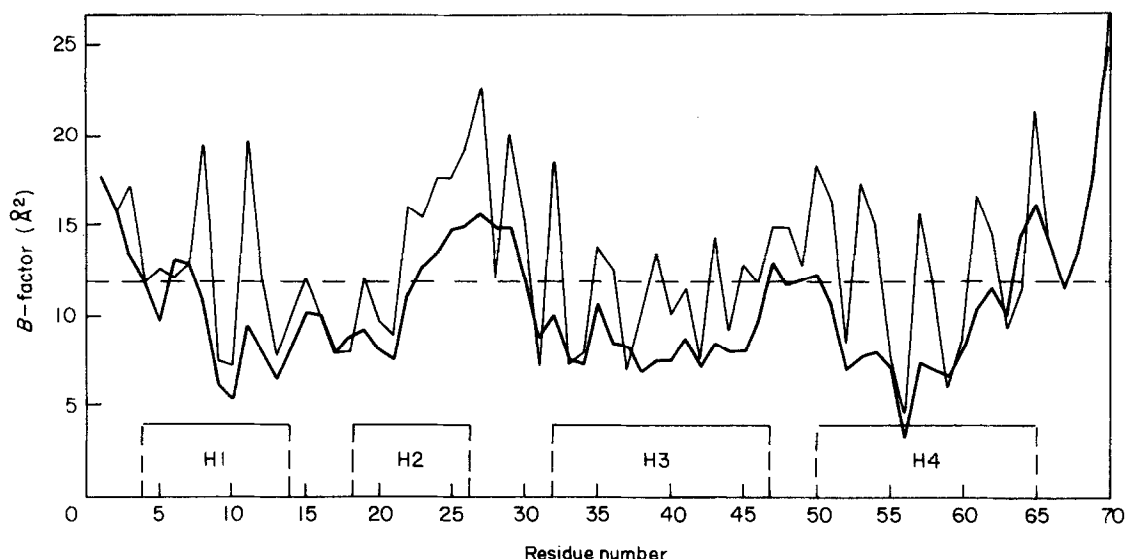


Figure 5. Mean temperature factor of the main chain atoms (thick line) and side-chain atoms (thin line) as a function of the residue number. The 4 α -helices of the monomer are marked and 11.9 Å² is the overall mean value.

The worst region of the final electron density map corresponds to the loop (Glu27–Phe28–Glu29), which appears to be a flexible part of uteroglobin.

The two disulfide bridges whose lengths were not restrained during the refinement, have the correct length of 2.0 Å, and are perfectly resolved in the map.

There was no evidence suggesting changes in the amino acid sequence. However, some initial assignments of atoms in amide amino acids were exchanged, relying essentially on their thermal factor. For example, NE2 and OE1 in Gln40, OD1 and ND2 in Asn12.

(b) Temperature factor

In the first cycle of the refinement we introduced an overall temperature factor estimated by Wilson's method with the program NORMAL (MULTAN) (Main *et al.*, 1977). Its value was 11.9 Å². Later on, individual thermal factors were refined.

A plot of the mean *B*-values for main-chain atoms (thick line) and side-chain atoms (thin line) in each amino acid is shown in Figure 5. The distribution is as expected: atoms of the main chain have lower *B*-values than those of the side-chains, and residues exposed to the solvent like Ile2, Lys26, Glu27, Glu29, Lys65 and Met70 have higher *B*-values (around 20 Å²) than internal residues like Val9, Ile10, Leu13, Ile56, Leu59 (*B* around 7 Å²).

Residues between helices are generally less well ordered than the central part of the helices, except for the loop between helix 1 and helix 2, which is stabilized by a hydrogen bond between N of Lys42 and O of Gly16 (length 2.80 Å).

Residues lining the central cavity generally have low *B*-values. For example, Ile56, Ile10, Leu59, Val19, Gly38, Leu13 are the six residues with the smallest thermal factors (less than 7.1 Å²). The highest average *B*-value is 25.9 Å² for Met70.

Since we introduced theoretical hydrogen atoms near the end of the refinement, the temperature factors of non-hydrogen atoms did not have to compensate for the lack of hydrogen atoms.

5. The Uteroglobin Structure

(a) Overall structure

The protein is a dimer and the assembly of its two subunits forms a thick shell leaving a cavity inside the molecule. The protein is shown in Figure 6 with its hydrogen-bonding interactions, which define its secondary structure. It is pear-shaped, with only one domain. Its size is about 41 Å × 35 Å × 31 Å along the principal axes *a*, *b* and *c*. The "pear" has a width of 31 Å at the "bottom", while it is 23 Å wide at the "top".

In spite of the low content of solvent, direct contacts between protein molecules are not numerous (cf. Table 10), while indirect contacts involving binding water molecules play a major role.

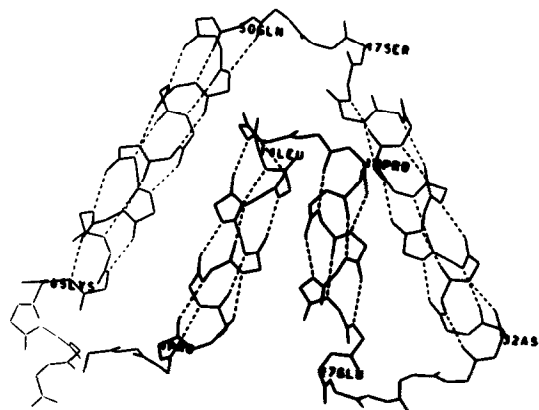


Figure 6. Hydrogen bonds CO_(i)...HN_(i+4) in the helices of the monomer. Residues at the beginning and the end of the helices are labeled.

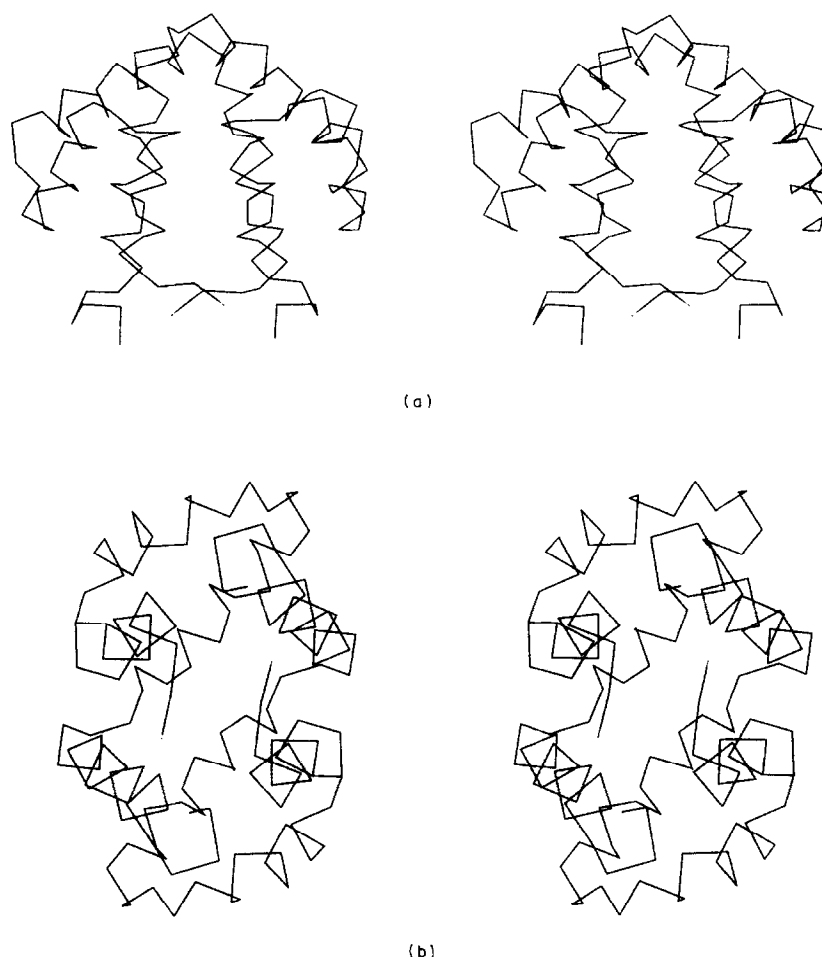


Figure 7. Stereo view of the C $^{\alpha}$ -chains of uteroglobin: (a) along the *c*-axis; (b) along the *b*-axis.

(b) Secondary structure

Uteroglobin is a protein consisting of α -helices throughout (see Fig. 7). There is no β -sheet and no long stretch of coil.

Each monomer starts with a small piece of irregular chain (from Gly1 to Cys3), followed by a proline residue (Pro4) that marks the beginning of the first helix (from Pro4 to Leu14, 11 residues). Then we encounter a small linear strand (Leu15, Gly16 and Thr17), which leads into the second helix (from Pro18 to Glu27, 10 residues). The only β -turn of the structure (Lys26, Glu27, Phe28 and Glu29) is located between helix 2 and helix 3.

At the N-terminal end of helix 3 (from Asp32 to Ser47, 16 residues) we find Pro30 and Asp31, and the two residues Leu48 and Pro49 link helix 3 to helix 4 (from Glu50 to Lys60, 16 residues). The remaining residues (from Ser66 to Met70) are folding without a particular secondary structure.

On the whole, 76% of the residues are part of α -helices and it is interesting to see the prolines dividing the secondary structure into segments.

By looking at the Ramachandran plots of φ , ψ we find these results again (cf. Fig. 8). In addition, the narrow distribution of the main-chain dihedral angle values concentrated around the mean values

of $\langle\varphi\rangle = -62^\circ$ and $\langle\psi\rangle = -42^\circ$ characteristic of α -helices, show their high level of regularity.

Two turns of the 3_{10} helix, Lys26, Glu27, Phe28 and Leu45, Asp46, Ser47, have been noted at the end of helix 2 and helix 3, respectively.

The helix turns are stabilized by hydrogen bonds between peptide amide and carbonyl groups of residues ($i, i+4$) for α -helix and ($i, i+3$) for the 3_{10} helix. Table 6 lists them and specifies hydrogen bonds between nearby side-chains in the same helix.

With $\varphi = 42^\circ$ and $\psi = 56^\circ$ the residue Glu29 is classified as a component of a reverse turn of type III.

(c) Tertiary structure

The eight helices of the uteroglobin dimer are essentially held together by hydrophobic forces, but also by some hydrogen bonds between side-chains. Table 6C gives an overview of them and Table 7 gives the angles between helices.

When we look at the C $^{\alpha}$ -chain of the monomer, there appears a "pseudobinary symmetry" that relates helix 1 to helix 2 on the one hand and helix 3 to helix 4 on the other (see Fig. 9). This could imply that the 4-helix chain originated from a 2-helix chain by gene duplication. However, the evidence

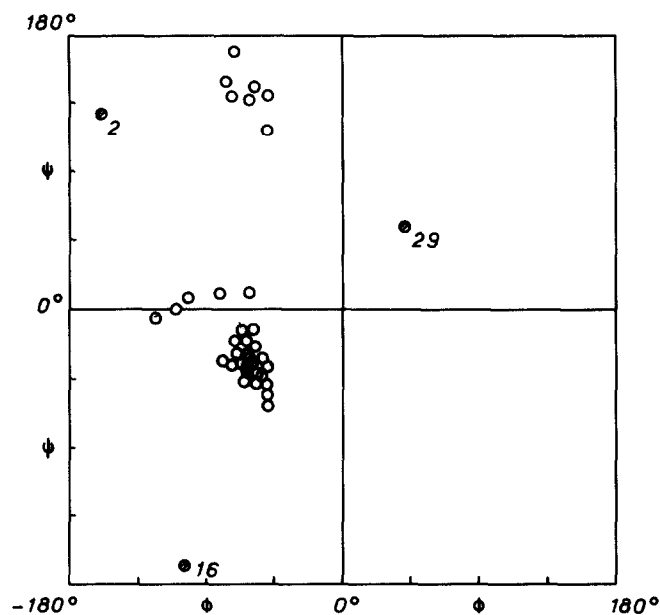


Figure 8. Ramachandran plot of ϕ and ψ angles.

in the amino acid sequence is really too vague to support this hypothesis. The symmetry is probably spurious.

(d) Solvent structure

Crystals of the $C222_1$ form of uteroglobin have a low content of solvent of only 35%, which involves

Table 6
Hydrogen bonds in the protein

A. α -Helical hydrogen bond parameters			
	N-H . . . O (Å)	N . . . O (Å)	N-H . . . O (°)
Mean value	2.08	2.97	153.47
r.m.s. dev.	0.14	0.12	9.62
Min. value	1.89	2.79	124.59
Max. value	2.44	3.27	172.46
B. Between side-chains in the same helix			
Donor	Acceptor	d (Å)	
T23 OG1	S20 O	2.65	
S24 N	T23 OG1	2.60	
K42 NZ	D46 OD2	2.69	
S47 OG	V44 O	2.69	
M34 N	D31 OD1	2.82	
D36 OD2	Q40 OE1	2.88†	
R53 NH1	P49 O	2.77	
T60 OG1	I56 O	2.83	
M70 OT	C69 O	2.71	
C. Between side-chains in different helices			
K35 NZ	E22 OE1	2.58	
K42 NZ	L13 O	2.54	
N12 ND2	S24 OG	2.77†	
K42 NZ	G16 O	2.80	
Y21 OH	G38 O	2.96	

† Or acceptor-donor.

Table 7
Angles between axis of helices

Helices	Angle (°)	π Angle (°)
1-2	151.6	28.3
1-3	49.4	
1-4	160.2	19.8
1-4'	147.0	33.0
2-3	150.7	29.3
2-4'	60.1	
3-3'	128.7	51.2
3-4	119.5	60.5
3-4'	107.8	72.2

1, 2, 3 and 4 are the numbers of the helices of one monomer (see Fig. 9) and 1', 2', 3' and 4' for the symmetrical helices in the second monomer. Angle is the value of the angle between the oriented axes of the 2 helices.

theoretically 185 water molecules in the asymmetric unit. During the refinement 87 peaks on the Fourier difference maps ($|F_o| - |F_c|$, φc) were assigned to water molecules; four of them were later excluded, since their parameters shifted to meaningless values. Therefore, 83 water molecules were retained (45% of the total solvent content), seven of them are located in the central cavity, and one in the bottom pocket.

All solvent peaks were refined as oxygen atoms but it is possible for some of them to be ammonium (NH_4^+) or sulfate (SO_4^{2-}) ions.

The temperature factors of these water molecules are in the range of 7.5 to 48 Å² and their mean value is 26 Å². The occupancy factors are in the range of 0.24 to 1.0, with 0.63 as the mean value.

The majority of water molecules are within the first hydration shell (<3.4 Å) and a number of them are directly hydrogen-bonded to the protein. The others are clustered with those of the second hydration shell giving a thick network of hydrogen bonds.

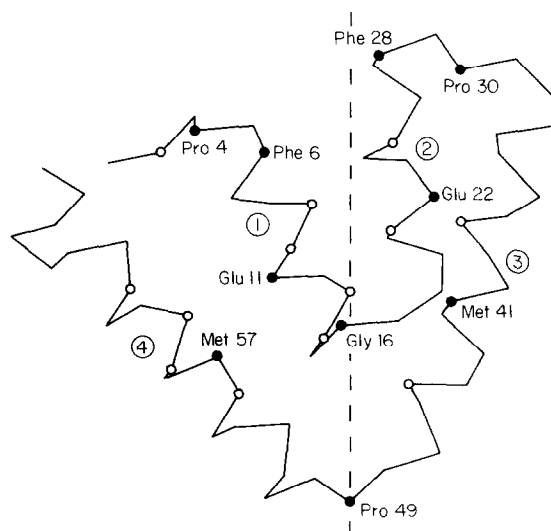


Figure 9. Pseudo symmetry within the uteroglobin monomer.

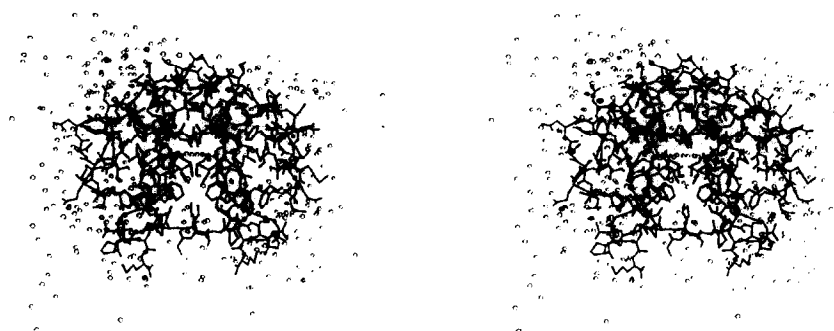


Figure 10. Stereo view of all the water molecules surrounding uteroglobin (along the *c*-axis).

A stereo view of all the water molecules surrounding the main chain of uteroglobin is shown in Figure 10.

Table 8 and Figure 11 summarize the hydrogen-bonded interactions and contacts involving water molecules. We can find some important values, such as: 71 contacts are with the protein and have an average length of 3.01 Å ($\sigma = 0.31$ Å) while 52 contacts are with other water molecules and have an average length of 2.81 Å ($\sigma = 0.36$ Å). Some of these distances are short, but this can be explained by the high *B*-values of the atoms concerned. For example, the smallest one (2.22 Å) is W49 ... O Pro30 with a *B*-value of 27 Å² for W49.

Between two water molecules the closest contacts may be explained, in part, by the occupancy factors of the considered sites because they may be occupied alternatively. For example, W3 and W69 are at the distance of 2.28 Å and their occupancies are 0.84 and 0.66, respectively.

Other contacts are between protein, water and another protein molecule but they are not numerous, even if they contribute greatly to the stability of the crystal (see Fig. 12).

The water molecules located in the central cavity are also hydrogen bonded to each other, as is discussed in the following section.

(e) Central cavity

Since the beginning of the X-ray study of uteroglobin, an oblong hydrophobic pocket has been shown to be present inside the protein (Mornon *et al.*, 1980).

This cavity is about the binary axis of the dimer and is lined by 24 hydrophobic residues (Ile2, Phe6, Val9, Ile10, Leu13, Leu14, Leu25, Met41, Leu45, Ile56, Leu59, Ile63) and six semi-hydrophilic ones (Tyr21, Gly38, Thr60), taking into account both subunits related by molecular symmetry.

In the first Fourier maps, at 2.2 Å resolution and without refinement, a very spread out volume of electron density led us to think that progesterone, or a similar steroid, could be present in oxidized uteroglobin. On the basis of the present structural analysis, this electron density may be interpreted as being due to the presence of 14 not very well ordered water molecules (7 in the asymmetric unit).

This interpretation was confirmed during the subsequent refinement and is in good agreement with the fading of electron density during the addition of high-resolution reflections.

These 14 water molecules (cf. Fig. 13) are held together by hydrogen bonds and four of them (W31, W31', W61, W61') are very close to protein

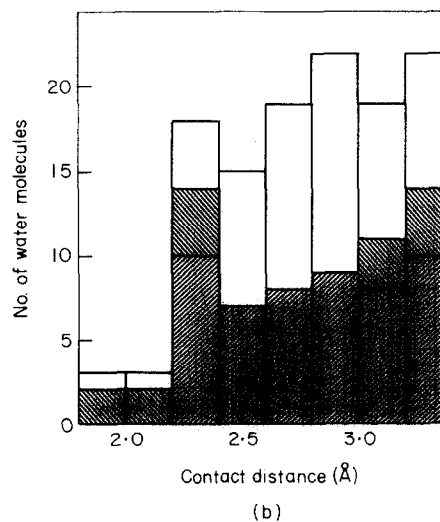
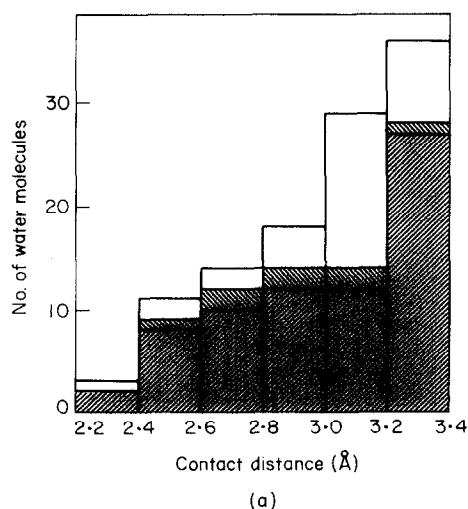


Figure 11. Histogram of contacts between water molecules and (a) protein atoms or (b) other water molecules. (▨) Contacts involving the water molecules of the asymmetric unit (*X*, *Y*, *Z*); (■) contacts involving the water molecules generated by the binary molecular axis (1-*X*, *Y*, 1-5-*Z*); (□) contacts involving the other water molecules.

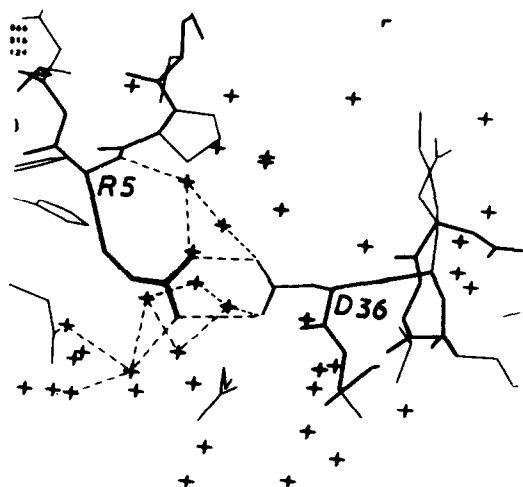


Figure 12. Examples of contacts between 2 protein and several water molecules. We notice direct contacts between protein residues (Arg5–Asp36) and also indirect contacts through water molecules.

atoms (OG1 Thr60', OG1 Thr60, OH Tyr21 and OH Tyr21', respectively, see Table 8C). In spite of these contacts the water molecules are not well ordered as is indicated by some very close contacts, water–water, and the appearance of the electron density. This disorder is probably the consequence of the hydrophobic nature of the cavity.

6. Discussion

(a) Central cavity

In the light of these results we propose a model of the interaction between progesterone and uteroglobin.

A calculation was done of the molecular surface (Vaney *et al.*, 1985) inside the protein as well as outside the protein (Fig. 14).

Figure 15 shows the surface of the central cavity and the van der Waals' surface of progesterone as

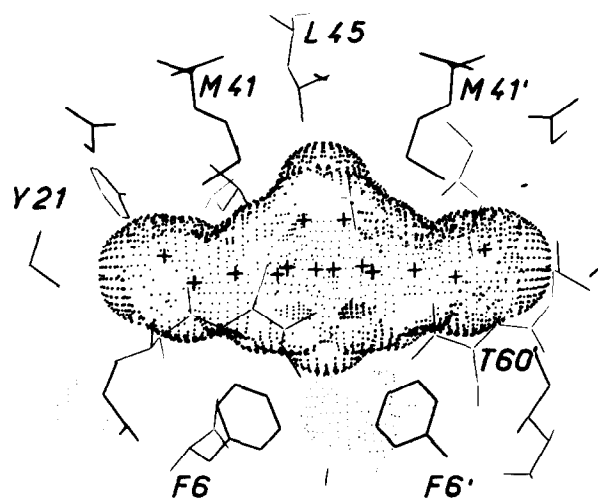


Figure 13. The 14 water molecules in the central cavity.

Table 8
Contacts involving water molecules (<3.40 Å)

A. With the protein					
Type	Number	Mean value	r.m.s. dev.	Min.	Max.
A.U.†	71	3.011	0.3155	2.220	3.390
Dim.	8	2.897	0.2729	2.470	3.390
Sym.	111	3.010	0.2972	2.220	3.390
Total	190	3.006	0.3030	2.220	3.390
B. With water molecules					
A.U.	52	2.806	0.3593	2.230	3.370
Dim.	15	2.613	0.6336	1.220	3.370
Sym.	54	2.802	0.3562	1.590	3.360
Total	121	2.780	0.391	1.220	3.370
C. In the central cavity					
W13:	W16 2.36	W31 2.26	W61 2.85	W72 3.09	
	W60 3.33				
W16:	W13 2.36	W60 2.42	W70' 2.31	W60' 2.14	
W31:	W13 2.26	W61 2.26	T60' OG1 2.66		
W60:	W70 2.29	W72 2.23	W16 2.42	W13 3.33	
	W16' 2.14	W70' 2.29	W60' 1.22		
W61:	W13 2.85	W31 2.26	Y21 OH 3.32		
W70:	W60 2.29	W16' 2.31	W70' 1.90	W60' 2.29	
W72:	W60 2.23	W13 3.09	W16' 3.37		
W72:	I10 CG1 3.04		L13 CD1 2.97		

† A.U., water of the asymmetric unit; Dim., water generated with 1–X, Y, 1.5–Z (dimer); Sym., water generated with other symmetry operations.

well as its backbone. The position of the progesterone molecule inside the cavity was obtained by the best visual superimposition of the two surfaces using a PS300 picture system and the software MANOSK (Vaney *et al.*, 1985).

A calculation of the van der Waals' signature (Surcouf, 1982; Surcouf & Mornon, 1982) of uteroglobin and progesterone was also performed in order to obtain more information about high-energy areas around the molecules.

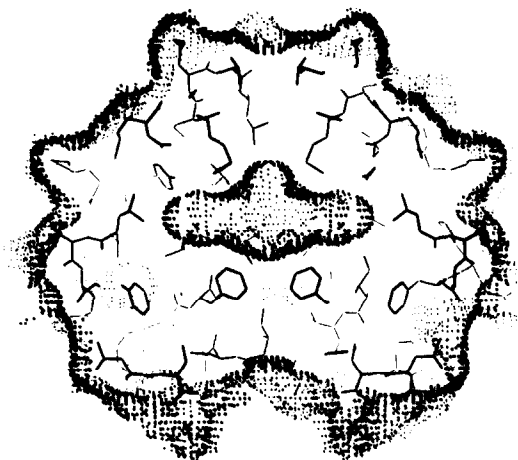


Figure 14. Clipped view along the *c*-axis, of the molecular surface of uteroglobin showing the central cavity and the pocket in the bottom part of the protein.

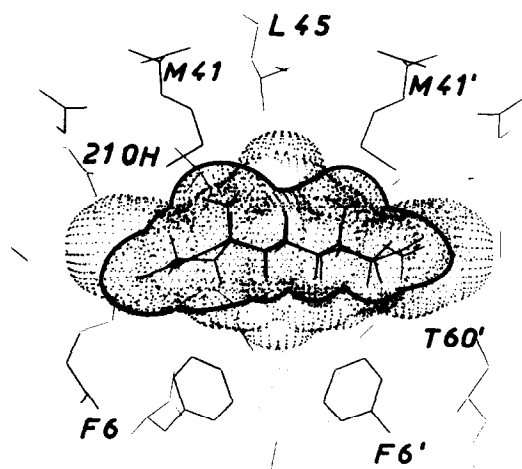


Figure 15. Simulation of the interaction between progesterone and uteroglobin.

It appears that the two surfaces match quite well and that progesterone may be hydrogen bonded to OH of Tyr21 and Tyr21' through O3 and O20, after small displacements of the two tyrosine residues. This superposition also suggests some displacement of the end of the two side-chains Met41 and Met41', to allow the methyl groups of progesterone to be accommodated completely by uteroglobin†.

(b) Surface pockets

The molecular surface calculation at the end of refinement led to surprising results. As expected, we observed the central cavity, but in addition four surface pockets related by symmetry were found.

Figure 16 shows the relative positions of central cavity (C), surface pockets (H1H4, H1H4') and the bottom pocket (P). The other two surface pockets (H1H2, H1H2') are not visible in this Figure. The names are derived from the positions: H1H4 is located between helix 1 and helix 4, and H1H2 between helix 1 and helix 2, the other two are symmetry-related.

Table 9 indicates the environment of each surface pocket. The most important residues are probably His8, Ile10 and Thr60. Ile10 and Thr60 are in the bottom of the H1H4 pocket but also belong to the central cavity. His8 is located between H1H4 and H1H2 pockets. Otherwise, Thr60 is the last residue coded by the main exon of uteroglobin (Bailey *et al.*, 1983). However, the significance of this finding is uncertain. His8 has been implemented in

† A study of the interaction between progesterone and uteroglobin is underway using molecular dynamics and two-dimensional nuclear magnetic resonance experiments. To reach the central cavity, the best access for the steroid seems to be through the "low" part of the protein, near the C and N termini. This process requires the reduction of the 2 disulfide bridges, then the moving of residues 1 to 6, together with some rotation of helices 1 and 4 and of their symmetrical counterparts.

Table 9
Residues environment

Central cavity					
6Phe	9Val	10Ile	13Leu	14Leu	21Tyr
25Leu	38Gly	41Met	56Ile	59Leu	60Thr
H1H2					
5Arg	8His	9Val	24Ser	26Lys	27Glu
28Phe					
H1H4					
1Gly	2Ile	7Ala	8His	10Ile	11Glu
57Met	60Thr	61Gln	64Val	65Lys	

progesterone binding (Lopez de Haro & Nieto, 1985).

The pockets H1H4 and H1H4' have essentially the shape of cylindrical pits that look like those situated near the catalytic site of serine proteases, and which generally accommodate lysine or arginine residues (Ruhlmann *et al.*, 1973). The pockets H1H2 and H1H2' are smaller but have a similar geometry. It seems unlikely that the pocket H1H4 alone could accommodate progesterone, but it could play a role in providing intermediate attachment sites during ligand binding. For example, the association of two uteroglobin dimers, each facing two H1H4 pockets, is conducive to good surface matching and creates an interstitial cavity that could readily accommodate a progesterone molecule (Vaney, 1986). This hypothetical association could explain the binding ratio 2 uteroglobin/1 progesterone, observed in experiments *in vitro* (Fridlansky & Milgrom, 1976).

In addition to being a potential carrier/protector of progesterone, it was suggested that uteroglobin, together with transglutaminase, may play a specific role in masking the antigenicity of developing embryos during implantation (A. B. Mukherjee *et al.*, 1980, 1982; D. C. Mukherjee *et al.*, 1980; Ramanathapuram *et al.*, 1984). In this mechanism,

Table 10
Intermolecular contacts ($< 3.0 \text{ \AA}$)

Donor	Acceptor	$d \text{ (\AA)}$	Symmetry
N12 OD1	T33 OG1†	(3.30)	1.5-X, 0.5-Y, -0.5+Z
T17 N	D31 OD2	2.83	
T17 OG1	D31 OD2†	2.71	
S20 OG	D31 OD1†	3.21	
S20 OG	D31 OD2†	2.56	
K58 NZ	S19 O	2.92	0.5+X, 0.5-Y, 1.0-Z
K58 NZ	E22 OE2	2.74	
K26 NZ	E54 OE1	2.77	
E22 OE1	M70 O	3.14	-0.5+X, 0.5+Y, Z
K35 NZ	M70 O	2.73	
R5 NH1	D36 OD1	2.57	1.5-X, 0.5+Y, 1.5-Z
R5 NH2	D36 OD2	3.16	
T17 CG2	K62 NZ	3.14	0.5+X, 0.5-Y, 1.0-Z
T23 CG2	N55 OD1	2.69	
T23 CG2	K58 NZ	2.68	

† Or Acceptor-Donor.

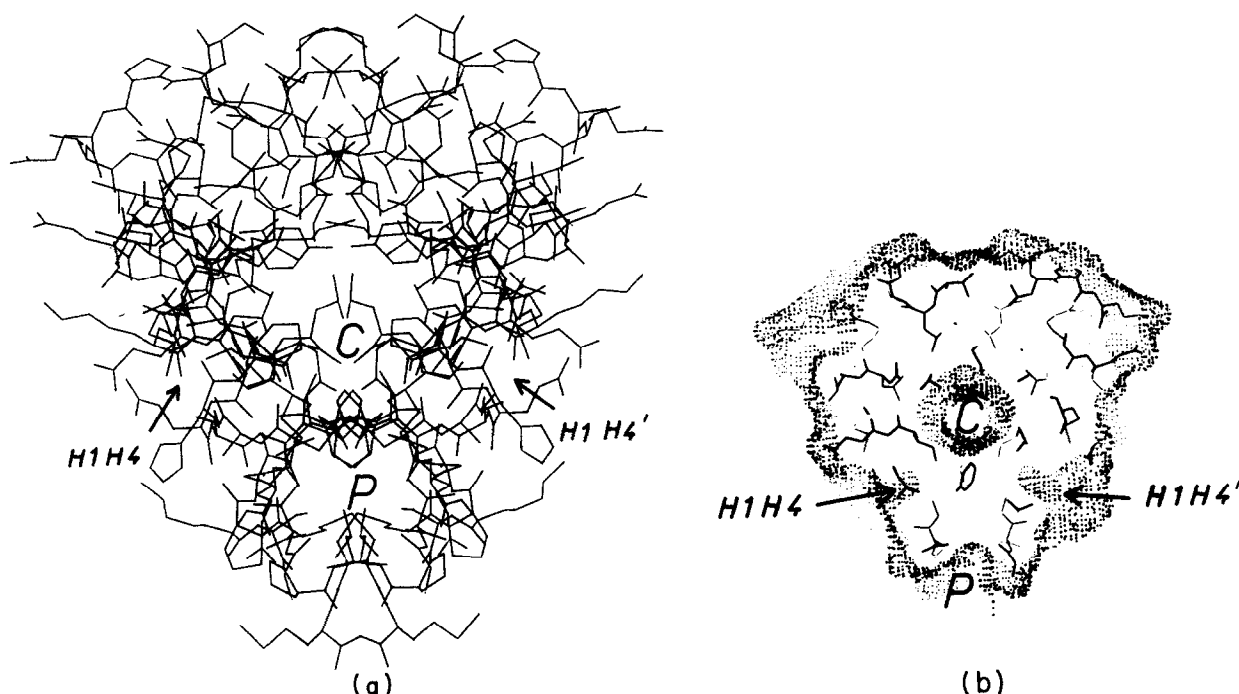


Figure 16. Clipped view along the *a*-axis showing the relative positions of the central cavity (C), the 2 surface pockets (H1H4 and H1H4') and the bottom pocket (P).

it was proposed that uteroglobin may cross-link with β 2-microglobulin on the embryonic cell surface. However, the structural arguments proposed to support this putative behavior rest on faulty data. The sequence of uteroglobin between residues 34 and 41, as quoted by Ramanathapuram *et al.* (1984) (M-K-D-A-G-Q-M-K), reported to be complementary to that of β 2-microglobulin between residues 1 and 8 (I-Q-R-T-P-K-I-Q) is wrong. The true sequence for uteroglobin in this region is (M-K-D-A-G-M-Q-M), so it leaves little chance for very specific cross-linking. However, the fact is that, with exception of Lys42, all lysine and glutamine residues of uteroglobin are freely accessible to cross-linking. A relationship between the H1H4 pocket and this possible immunosuppressive function of uteroglobin may exist, but it is entirely speculative.

Until now, we have calculated the surfaces of only a relatively small number of proteins and have not yet attempted a systematic comparison of their characteristics. So it is too early to speculate further about the significance of the similarity between H1H4 surface pockets of uteroglobin and the pits of serine proteases that are situated near their catalytic pockets. Of course, it may be only the result of chance or, alternatively, it will be followed by similar observations suggesting specific convergent evolution between surface structures and functions for different classes of proteins. It is of particular interest to note that a completely independent hypothesis was proposed by Backer (1984) and Lukola & Punnonen (1982), suggesting the presence of serine protease-like binding pockets

on uteroglobin and other steroid binding proteins. The present observations strongly support this.

From the few complete surface calculations we have made, another striking general structural homology was discovered between uteroglobin and phospholipase A₂ from snake venom.

Phospholipase A₂ (Brunie *et al.*, 1985) is also built from two chemically identical monomers that, in the crystalline form studied, are related to each

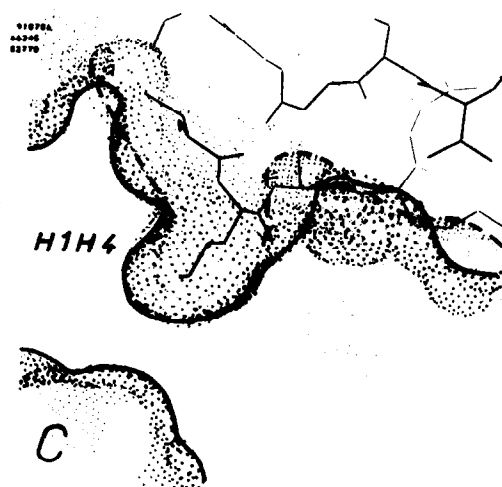


Figure 17. Superposition of the active site of trypsin and the H1H4 surface pocket of uteroglobin. Lys15 of bovine pancreatic trypsin inhibitor (BPTI) in the active site of trypsin is shown (---, trypsin; —, uteroglobin).

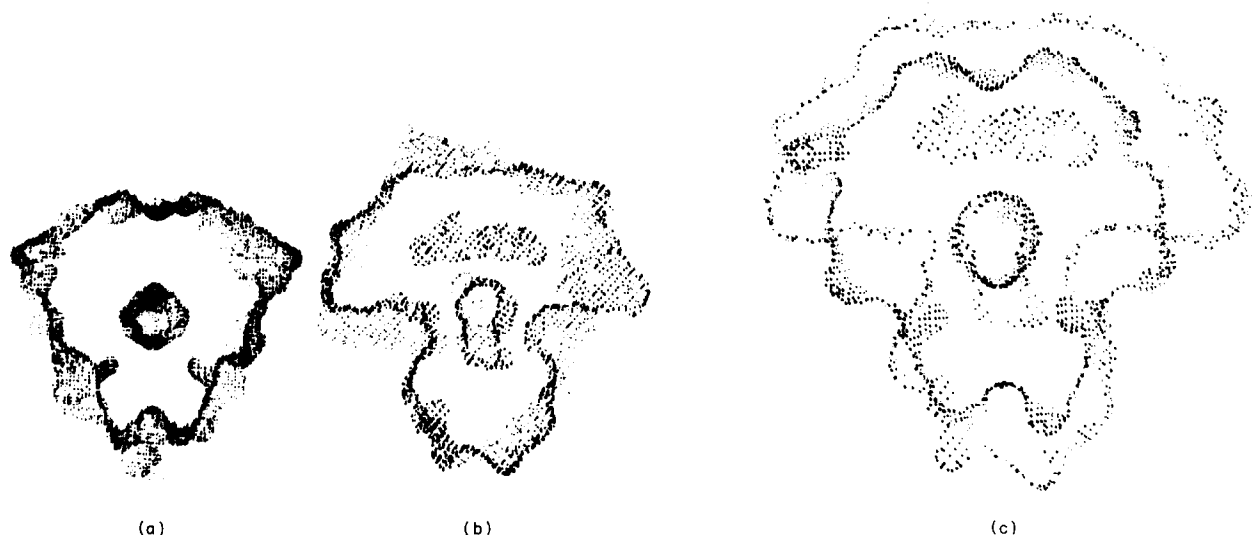


Figure 18. Comparison of phospholipase and uteroglobin surfaces. Clipped views showing the similarities between the 2 proteins. (a) Uteroglobin; (b) phospholipase; (c) superposition.

other by a pseudo binary axis, as uteroglobin is in its $P1$ or $P2_1$ crystal forms.

Phospholipase A_2 possesses two internal cavities whose size is similar to the uteroglobin cavities, and which are arranged perpendicularly to each other. In addition, two large surface pockets are present near one of the internal cavities. The relative positions of these features, in uteroglobin and in phospholipase A_2 are surprisingly similar (Fig. 18). As in uteroglobin, where Thr60 connects the central cavity and the surface pocket H1H4, Lys60 connects the central cavity and the surface pocket in phospholipase A_2 .

A detailed study of these homologies is underway and will be reported elsewhere.

We thank Dr S. Brunie for the stimulating discussion and for her permission to use the refined co-ordinates of phospholipase A_2 before publication.

References

- Atger, M., Mercier, J. C., Haze, G., Fridlansky, F. & Milgrom, E. (1979). *Biochem. J.* **177**, 985–988.
- Atger, M., Mornon, J. P., Savouret, J. F., Loosfelt, H., Fridlansky, F. & Milgrom, E. (1980). In *Steroid Induced Uterine Proteins* (Beato, M., ed.), pp. 341–350, Elsevier/North-Holland, Amsterdam and New York.
- Backer, M. E. (1984). *FEBS Letters*, **175**, 41–44.
- Bailly, A., Atger, M., Atger, P., Cerbon, M. A., Alizon, M., Mai Thu Vu Hai, Logeat, F. & Milgrom, E. (1983). *J. Biol. Chem.* **258** (17), 10384–10389.
- Beato, M., Arnemann, J., Heins, B., Muller, H. & Nieto, A. (1980). In *Steroid Induced Uterine Proteins* (Beato, M., ed.), pp. 351–368, Elsevier/North-Holland, Amsterdam and New York.
- Beier, H. M. (1968). *Biochim. Biophys. Acta*, **160**, 289–291.
- Brunie, S., Bolin, J., Getwinth, D. & Sigler, P. B. (1985). *J. Biol. Chem.* **260** (17), 9742–9749.
- Buehner, M. & Beato, M. (1978). *J. Mol. Biol.* **120**, 337–341.
- Buehner, M., Lifchitz, A., Bally, R. & Mornon, J. P. (1982). *J. Mol. Biol.* **159**, 353–358.
- Bullock, D. W. (1977). *Biol. Reprod.* **17**, 104–107.
- Bullock, D. W., Woo, S. L. C. & O'Malley, B. W. (1976). *Biol. Reprod.* **15**, 435–443.
- Fitzwater, S. & Scheraga, H. (1980). *Acta Crystallogr. sect. A*, **36**, 211–219.
- Fridlansky, F. & Milgrom, E. (1976). *Endocrinology*, **99**, 1244–1251.
- Furey, W., Wang, B. C. & Sax, M. (1982). *J. Appl. Crystallogr.* **15**, 160–166.
- Hendrickson, W. A. & Konnert, J. H. (1980a). In *Computing in Crystallography* (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp. 13.01–13.23, Indian Institute of Science, Bangalore.
- Hendrickson, W. A. & Konnert, J. H. (1980b). In *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., ed.), vol. 1, pp. 43–47, Pergamon, Oxford.
- Hermans, J. & McQueen, J. E. (1974). *Acta Crystallogr. sect. A*, **30**, 730–739.
- Jack, A. & Levitt, M. (1978). *Acta Crystallogr. sect. A*, **34**, 931–935.
- Jones, T. A. (1982). In *Computational Crystallography* (Sayre, D., ed.), pp. 303–317, Clarendon Press, Oxford.
- Lopez de Haro, S. & Nieto, A. (1983). *Arch. Biochem. Biophys.* **266** (2), 539–547.
- Lopez de Haro, S. & Nieto, A. (1985). *FEBS Letters*, **193** (12), 247–249.
- Lukola, A., Punnonen, M. (1982). *Biochem. Biophys. Res. Commun.* **108**, 822–827.
- Luzatti, V. (1952). *Acta crystallogr.* **5**, 802–810.
- Main, P., Lessinger, L., Woolfson, M. M., Germain, G. & Declercq, J. P. (1977). *Multan. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*, University of York, UK, and University of Louvain, Belgium.
- Mornon, J. P., Fridlansky, F., Surcouf, E., Bally, R. & Milgrom, E. (1978). *J. Mol. Biol.* **122**, 237–239.
- Mornon, J. P., Fridlansky, F., Bally, R. & Milgrom, E. (1979). *J. Mol. Biol.* **127**, 237–239.
- Mornon, J. P., Fridlansky, F., Bally, R. & Milgrom, E. (1980). *J. Mol. Biol.* **137**, 415–429.

- Mukherjee, A. B., Laki, K. & Agrawal, A. K. (1980). *Science*, **219**, 989–991.
- Mukherjee, A. B., Ulane, R. E. & Agrawal, A. K. (1982). *Amer. J. Reprod. Immunol.* **2**, 135–141.
- Mukherjee, D. C., Agrawal, A. K., Ramanathrapuram, M. & Mukherjee, A. B. (1980). *Med. Hypothesis*, **6**, 1043–1055.
- Noske, I. & Feigleson, M. (1976). *Biol. Reprod.* **15**, 704–713.
- Ponstingl, H., Nieto, A. & Beato, M. (1978). *Biochemistry*, **17**, 3908–3912.
- Popp, R. A., Foresman, K. R., Wise, L. D. & Daniel, J. C., Jr (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 5516–5519.
- Ramanathrapuram, M., Soo It Chung & Mukherjee, A. B. (1984). *Biochem. Biophys. Res. Commun.* **121** (1), 400–407.
- Ruhlmann, A., Kukla, D., Schwager, P., Bartels, K. & Huber, R. (1973). *J. Mol. Biol.* **77**, 417–436.
- Savouret, J. F. & Milgrom, E. (1983). *DNA*, vol. 2, pp. 99–104, M. A. Liebert, Inc. Publ.
- Surcouf, E. (1982). *Contribution à l'étude des interactions "protéines réceptrices/substrats". Exemples d'application Cas particulier du récepteur progestogène*, thèse de doctorat d'état, Université Paris VI.
- Surcouf, E. & Mornon, J. P. (1982). *C. R. Acad. Sci., Paris*, ser. 2, **295** (11), 923–926.
- Tancredi, T., Temussi, P. A. & Beato, M. (1982). *Eur. J. Biochem.* **122**, 101–104.
- Tuula, T. (1980). *Biochem. Biophys. Res. Commun.*, **97** (2), 559–565.
- Vaney, M. C. (1986). *Conception d'un logiciel interactif de graphisme et de simulation moléculaire: logiciel MANOSK. Exemples d'application: Modélisation de la structure tridimensionnelle de la "Prostatic Binding Protein" (PBP)*, thèse de doctorat, Université Paris VI.
- Vaney, M. C., Surcouf, E., Morize, I., Cherfils, J. & Mornon, J. P. (1985). *J. Mol. Graphics*, **3** (3), 123–124.

Note added in proof. The suggested displacements of Met41 and Met45, allowing the putative binding of progesterone (see section 6(a)), are actually observed in the refined P2 form of uteroglobin R. Bally, personal communication).

Edited by R. Huber