Journal of Computer-Aided Molecular Design, **16:** 917–934, 2002.
KLUWER/ESCOM

On the molecular interaction between lactoferrin and the dye Red HE-3B. A novel approach for docking a charged and highly flexible molecule to protein surfaces

Mariano Grasselli[1,4], Osvaldo Cascone[1], F. Birger Anspach[3,5] & Jose M. Delfino[2,*]
[1]*Cátedra de Microbiología Industrial y Biotecnología and* [2]*Departamento de Química Biológica-IQUIFIB, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956, 1113 Buenos Aires, Argentina;* [3]*GBF-Gesellschaft für Biotechnologische Forschung, Biochemical Engineering Division, Mascheroder Weg 1, D-38124 Braunschweig, Germany* [4]*Present address: Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Roque Sáenz Peña 180. (B1876BXD) Bernal, Provincia de Buenos Aires, Argentina;* [5]*Present address: Hochschule für Angewandte Wissenschaften Hamburg, Fachbereich Naturwissenschaftliche Technik, Lohbrügger Kirchstrasse 65, D-21033 Hamburg, Germany*

Key words: conformational analysis; dye-protein complex; electrostatic interaction; lactoferrin; molecular modeling; protection from proteolysis; Red HE-3B; docking; flexible ligand

Summary

Lactoferrin (Lf) is a non-heme, iron binding protein present in many physiological fluids of vertebrates where its main role is the microbicidal activity. It has been isolated by different methods, including dye-affinity chromatography. Red HE-3B is one of the most common triazinic dyes applied in protein purification, but scant knowledge is available on structural details and on the energetics of its interaction with proteins. In this work we present a computational approach useful for identifying possible binding sites for Red HE-3B in apo and holo forms of Lfs from human and bovine source. A new geometrical description of Red HE-3B is introduced which greatly simplifies the conformational analysis. This approach proved to be of particular advantage for addressing conformational ensembles of highly flexible molecules. Predictions from this analysis were correlated with experimentally observed dye-binding sites, as mapped by protection from proteolysis in Red HE-3B/Lf complexes. This method could bear relevance for the screening of possible dye-binding sites in proteins whose structure is known and as a potential tool for the design of engineered protein variants which could be purified by dye-affinity chromatography.

Abbreviations: bLf, bovine lactoferrin; hLf, human lactoferrin.

Lactoferrin (Lf) is a non-heme, iron binding protein, originally isolated from milk [1]. This molecule has been identified in a variety of physiological fluids of vertebrates [2] and it is also a major component of specific granules of neutrophilic leukocytes [3]. Biological properties ascribed to this protein include the regulation of absorption of iron (and other metal ions) in the gastrointestinal tract, and the modulation of both the production of polymorphonuclear leukocytes and animal cell growth. Lf also shows microbicidal activity against bacteria and yeast [3]. More recently, several authors have found that Lf can directly damage the Gram-negative bacterial outer membrane: Lf causes the release of lipopolysaccharide molecules from the membrane, thus enhancing the bacterial susceptibility to hydrophobic antibiotics and to human lysozyme [4, 5].

Lf is a monomeric glycoprotein of approximately 80 kDa molecular weight, constituted by two major lobes, each possessing one site capable of reversibly binding a ferric ion concomitantly with a carbonate anion. This cooperativity between metal and anion

*To whom correspondence should be addressed. E-mail: rtdelfin@criba.edu.ar

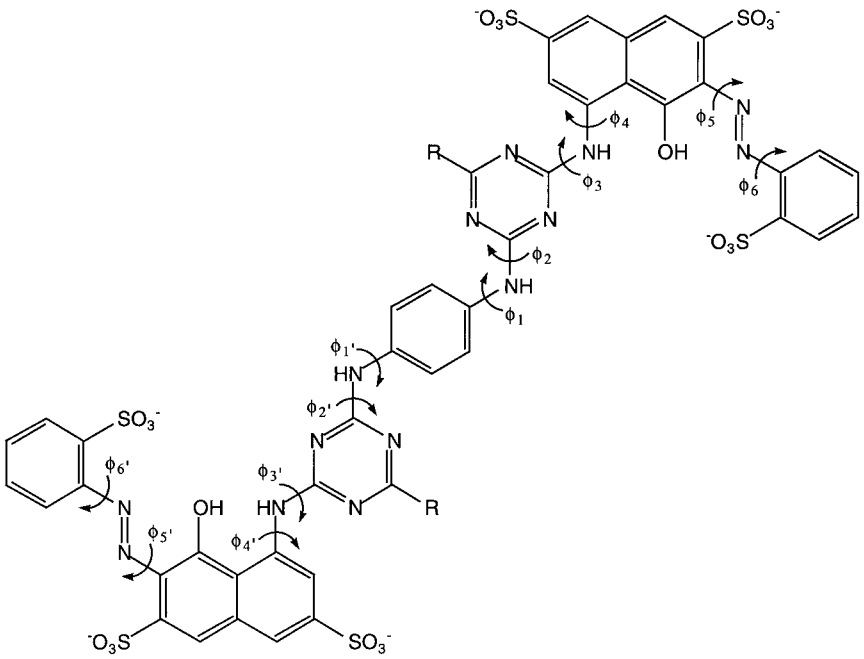


Figure 1. Structure of Red HE-3B showing the torsion angles determining the conformation. The substituents represented by R correspond to either OH or Cl.

binding is a characteristic property of transferrins [6]. Crystallographic structures of Lf have been determined, including the apo and holo forms of human Lf (hLf) [7, 8] and the holo form of bovine Lf (bLf) [9]. The former studies revealed that a relative movement of subdomains occurs upon iron binding [10].

Lf has been isolated from bovine whey by different chromatographic methods, including hydrophobic interaction, affinity and ion-exchange chromatography [11–15] , the latter being the method of choice in large-scale processes. Nevertheless, affinity methods are expected to show a better performance than ion-exchange procedures on account of the very low concentration of Lf present in bovine whey [16, 17]. In this regard, triazinic dyes, such as Cibacron Blue F3G-A and Red HE-3B, have been extensively employed to prepare affinity chromatography matrices due to their low cost, ready availability, simplicity of the immobilization reaction and resistance to biological and chemical degradation [17]. All these advantages make these compounds highly eligible for industrial scale protein purification [19]. Recently, we found that it is possible to recover bLf in high yield from whey using Red HE-3B-Sepharose under non-denaturing conditions [17].

In order to gain further insight into the nature of this dye-protein interaction, in this work we undertook the task of investigating key structural features in both Red HE-3B and Lf responsible for this phenomenon. It is well known that polysulfonated aromatic chromophores of triazinic dyes mimic the binding of naturally ocurring nucleotide coenzymes such as NAD^+ or $NADP^+$ to many enzymes [20]. Support for this evidence was substantiated by the crystallographic structure of lactate dehydrogenase-Cibacron Blue F3G-A complex [21]. Despite this fact, many other proteins including Lf, which do not associate with these coenzymes, also exhibit strong interaction with triazinic dyes [22]. The interaction of Lf with triazinic dyes was investigated by zonal chromatography [23] and differential circular dichroism [24]. With the latter technique, a distinct behaviour was observed when hLf or bLf were titrated with Cibacron Blue F3G-A. A different number of interaction sites as well as a different interaction mechanism were proposed for hLf and bLf apo and holo forms. In this regard, differential visible spectroscopy was advanced as a suitable method for studying interactions of proteins and dyes [25]. Spectral changes provide important clues on the chemical nature of the protein-dye interaction [26] and allow the determination of the number of binding sites and the dissociation constant of the complexes [27]. In spite of this, no detailed structural information can be derived from these studies.

In this paper, the interaction of Red HE-3B with hLf and bLf in both their apo and holo forms was analyzed by molecular modeling. Since no structural data on the dye is available, in the first instance a conformational analysis of Red HE-3B was carried out to evaluate the most likely structures. In the context of the vast literature dealing with docking methods (for a recent review see reference [28]), intense effort is currently being devoted to effectively introduce flexibility in the search process. In this regard, Red HE-3B poses a challenge, since it is an extremely flexible molecule which binds to target sites in proteins mainly through electrostatic interactions. We advance here a methodology which effectively introduces a simplified description of the ligand molecule to derive relevant geometrical constraints from the conformational ensemble calculated for this dye. Next, target protein surfaces are exhaustively surveyed to search for potential binding sites which satisfy those geometrical restrictions. The screening method rests upon electrostatic pairing between the negatively charged dye and positively charged arrays on the surface of the proteins. Predictions from this analysis were tested experimentally against data on protection from enzymatic proteolysis occurring upon binding of the dye to Lf.

Table 1. Sites in Lfs protected from tryptic proteolysis[a]

Protein	Position	Sequence
apo hLf	122-133	TAGWNVP(IGTLR)[c]
	225-236	DEYELLCPDNTR
	436-442	LAVAVVR[b]
	458-465	SCHTAVDR
	534-546	CLAENAGDVFVK
	662-675	YLGPQYVA(GITNLK)[c]
holo hLf	31-38	GPPVSCLK
apo bLf	8-18	WCTISQPEWFK
	40-47	AFALECIR
	74-85	LRPVAAEIYGTK
	152-163	FFSASCVPCIDR
	175-186	GEGENQCACSSR
	190-197	FGYSGAFK[b]
	270-274	LLSK
	545-562	NDTVWENTNGE(STADWAK)[c]
holo bLf	8-18	WCTISQPEWFK
	40-47	AFALECIR
	74-85	LRPVAAEIYGTK
	190-197	FGYSGAFK[b]
	310-319	IPSKVDSALY
	314-319	VDSALY

[a]Samples of each indicated protein were digested with trypsin in the absence or in the presence of Red HE-3B and the peptide mixtures were analyzed by HPLC, as described in Materials and methods. Those peptides appearing in the former, but not in the latter are listed here.
[b]Peptides resulting from a chymotryptic cleavage.
[c]Sections between parentheses were not actually sequenced, but deduced from the known amino acid sequence of each protein.

Results and Discussion

Conformational analysis of Red HE-3B

Optimized Monte Carlo search and energy minimization

Red HE-3B is a symmetrical molecule where, in principle, twelve rotational degrees of freedom can be identified (Fig. 1). However, an extended conjugated system would pose constraints to torsion around these bonds. A planar molecular model for this dye, as depicted in Fig. 1, was initially created using the model-building facility implemented in MacroModel [29] In each half molecule, azo bonds were fixed in their *trans* configuration. Coordinates for this structure were used as input for BatchMin, the calculation module of this program. From an experimental standpoint, the chlorine atoms in the dye (located at the R positions) would have reacted with nucleophiles (e.g. from the chromatographic support) by the time the noncovalent interaction with the protein occurs. Thus, for the purpose of this computational analysis, the hydrolysis product of the dye was considered, i.e. hydroxyl groups would replace the reactive chlorine atoms at the R positions.

In order to search the conformational space available to this compound, we ran an optimized Monte Carlo (MC) protocol [30] that considered all twelve possible torsions (ϕ_1 to ϕ_6 and $\phi_{1'}$ to $\phi_{6'}$) which determine the spatial structure. Between ten and fifty thousand conformers of the dye were analyzed following this MC method. The energy of each conformer was calculated *in vacuo* with the MM2* force field, setting the extended non-bonded distance cut-off to 50 Å for van der Waals and electrostatic interactions. Structures (2261 conformers) ranging in energy up to 50 kJ/mol above the global minimum were selected. A similar calculation was run employing water as the solvent (according to the semianalytical solvation treatment GB/SA [31]) which yielded 4496 low energy conformers.

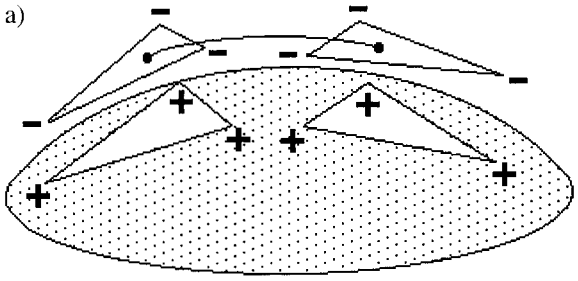


Figure 2a. Electrostatic model of the dye-protein interaction. Face-to-face matching between a pair of tethered triangles of negatively charged atoms (representing the dye) and a similarly shaped array of positively charged atoms occurring on the surface of the protein.

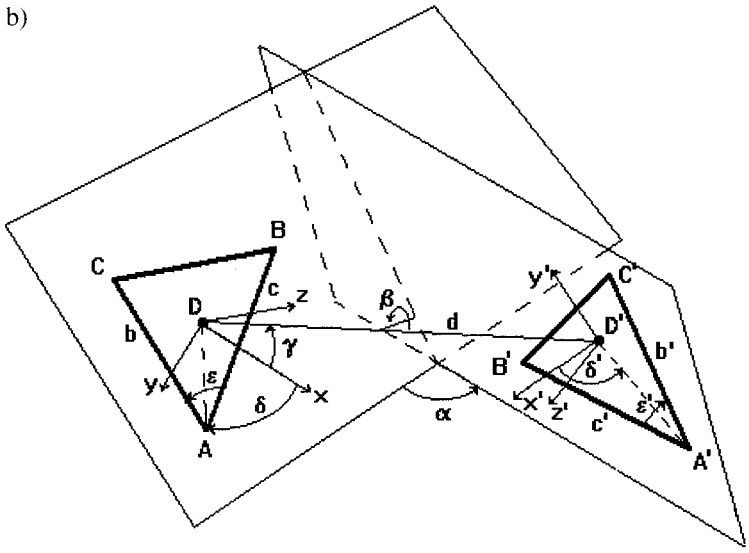


Figure 2b. Simplified geometrical description of Red HE-3B. The parameters shown correspond to those defined in that section of the main text.

The large number of conformers found for Red HE-3B reflects the extremely high flexibility of this molecule. This precludes further analysis based solely on the global minimum structure. Thus, in order to build a realistic molecular picture of the conformational ensemble for this molecule, we considered instead the statistical distribution of values for all torsion angles ϕ_i. Dihedral angles at both ends of each trans diazo bonds are restricted within defined intervals: ϕ_5 and $\phi_{5'}$ adopt values in the ranges 100° to 140° and −100° to −140°. Similarly, ϕ_6 and $\phi_{6'}$ fall within the intervals 100° to 150° and −100° to −150°. However, angles ϕ_5 (or $\phi_{5'}$) and ϕ_6 (or $\phi_{6'}$) are not independent from each other, as reflected by a cross correlation analysis: only two of four possible combinations are permitted (results not shown). These correspond to conformations where the sulfonic groups in ortho adopt opposite orientations to minimize steric hindrance and electrostatic repulsion. As a consequence of this fact, in both allowed conformations the terminal sulfonate lies at about the same distance to those substituting the naphthalene ring. In general, in each moiety of Red HE-3B the naphthalene ring and the terminal (2-sulfonyl)-phenyl group connected through the diazo link were found not to be coplanar, but to lie in roughly parallel planes.

For comparison, the distribution of values for the angles ϕ_5 and ϕ_6 was also studied for a hemimolecule of Red HE-3B, namely, the moiety substituting the central p-phenylene ring. This analysis yielded the following results: ϕ_5 lies in the ranges (70°,120°) and (−70°, −120°), while ϕ_6 lies in the ranges (110°, 130°) and (−110°, −130°). Here, the observed shifts in both angles as compared with the parent molecule might be due to the presence of long-range electrostatic repulsive interactions in the latter. In addition, hydrogen bonding was frequently observed between sulfonate groups on one side and the

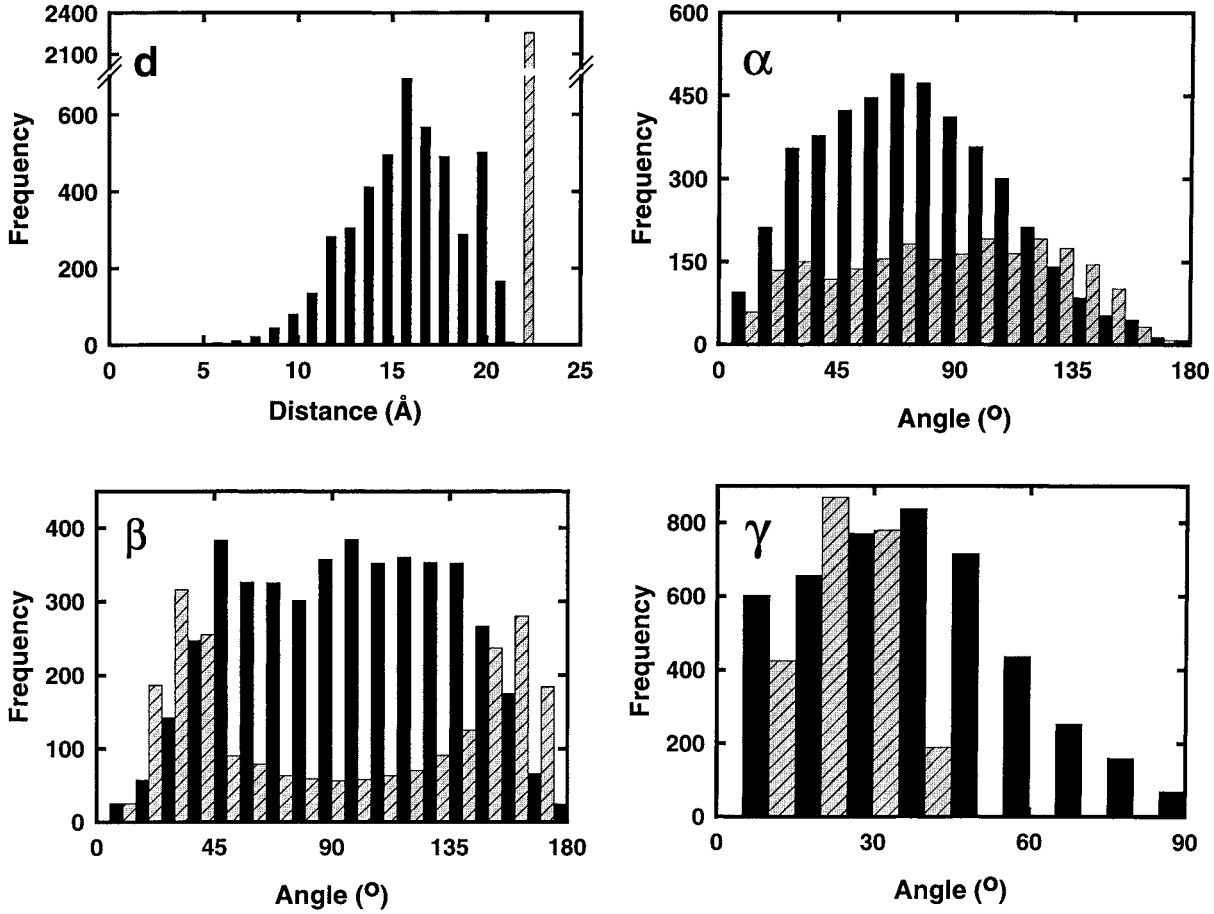


Figure 3. Frequency histograms resulting from the conformational analysis of Red HE-3B. The population of low energy conformers of the dye was described in terms of the parameters shown in Fig. 2b. The distributions of values for parameters d, α, β and γ are shown for simulations run in vacuo (dashed grey bars) and in water (black bars).

hydroxyl group substituent located on the opposite triazine ring. Similarly, the distributions of values for the other pairs of dihedral angles varied within narrow ranges: ϕ_1 (or $\phi_{1'}$): (−40°, −10°) and (10°,40°); ϕ_2 (or $\phi_{2'}$): (−10°,10°) and (170°, −170°); ϕ_3 (or $\phi_{3'}$): (−10°,10°) and (170°,−170°); and ϕ_4 (or $\phi_{4'}$): (−60°,−40°) and (40°,60°).

Simplified geometrical description of Red HE-3B

In order to achieve the main purpose of this work, which is the study of the interaction between Red HE-3B and target Lfs, we introduced a simplified geometrical description of the ligand to deal with the multiplicity of conformations found for this molecule. Previous evidence from differential UV-vis spectroscopy on dye-Lf complexes indicated that the nature of the interaction is largely electrostatic (Grasselli, M., Cascone O., Anspach F.B. unpubl.). Therefore, special emphasis was put on the spatial distribution of the negative charges at the sulfonate groups.

The dye molecule can be regarded as a pair of triangles (ABC and A′B′C′), the vertices of which corresponding to the sulfur atoms, joined by a relatively flexible tether (Fig. 2a). In this fashion, each triangle defines a plane and, upon these two planes, twelve parameters (and a pseudo parameter) were introduced to completely describe the internal geometry of the molecule (Fig. 2b). Six of these parameters define the geometrical relationship between the two triangles, namely, these are: d, the distance between the centers of mass of each triangle (D and D′); α, the dihedral angle between the planes defined by each triangle; β, the dihedral angle along the minimal distance (normal) between segment d and the line defined by the intersection of the two planes; γ (or γ′), the an-

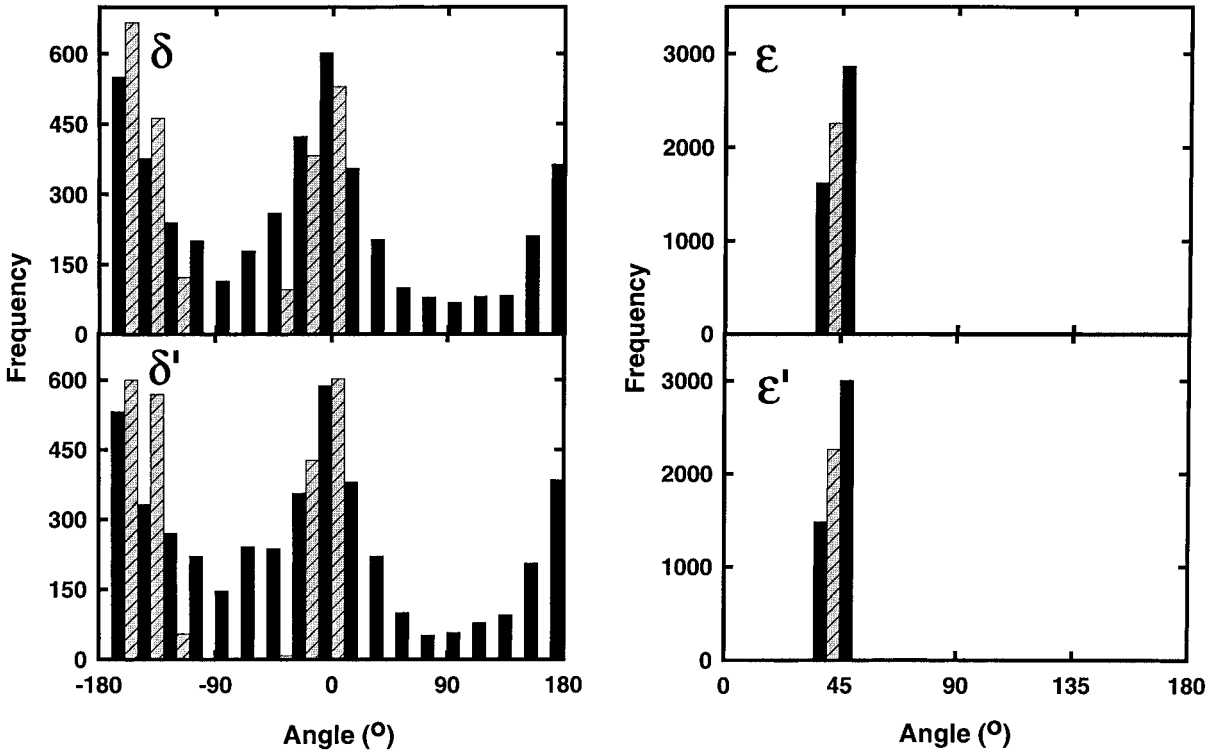


Figure 4. Frequency histograms resulting from the conformational analysis of Red HE-3B. The population of low energy conformers of the dye was described in terms of the parameters shown in Fig. 2b. The distributions of values for parameters δ, δ′, ε and ε′ are shown for simulations run in vacuo (dashed grey bars) and in water (black bars).

gle between d and its normal projection on each plane (which defines the x or x′ axes, respectively); and δ (or δ′), the angle between the x (or x′) axis and segment DA (or D′A′), where, for the purpose of this analysis (as will be explained below), A (or A′) corresponds to the position of the sulfur atom of the sulfonate group most proximal to the center of the molecule. Note that once γ has been defined, then γ′ becomes a pseudo (non-independent) parameter. In addition, another set of six parameters defines the internal geometry of each triangle, namely, side b (or b′) is the shorter segment including vertex A (or A′), and c (or c′) is the correspondingly longer segment; and ε (or ε′) is the angle between these segments. This set is equivalent to that defined by the internal coordinates ($3 \times 6 - 6 = 12$) of the sulfur atoms.

Analysis of the low energy conformers of Red HE-3B in terms of the new set of geometrical parameters

The minimum energy conformers found for Red HE-3B were parametrized according to the geometrical definition advanced in Fig. 2b and the frequency histograms for each parameter are shown in Figs. 3, 4 and 5.

For simulations run *in vacuo* the distribution of distances d is very sharp, with a maximum at 23 Å, a situation reflecting the maximal stretching suffered by this flexible molecule due to electrostatic repulsion between the sulfonated ends. By contrast, charge shielding in water causes a wide distribution of lengths (11–22 Å), where the prevailing conformers adopt a significantly less extended shape (d ∼16 Å). This latter situation is expected to represent more realistically the conformational ensemble adopted by Red HE-3B in aqueous solution, and these data will be used for further analysis.

The interplanar angle α shows a broad asymmetric distribution which extends between 0° and 160°, and which is centered at around 70° (in vacuo the distribution flattens to the point that any value in the range 20–140° becomes equally probable). Likewise, in water angle β shows a wide distribution, albeit more symmetrical than α between 10° and 170° (in vacuo a bimodal distribution is observed exhibiting maxima at 35° and 145°). In addition, in vacuo a cross-

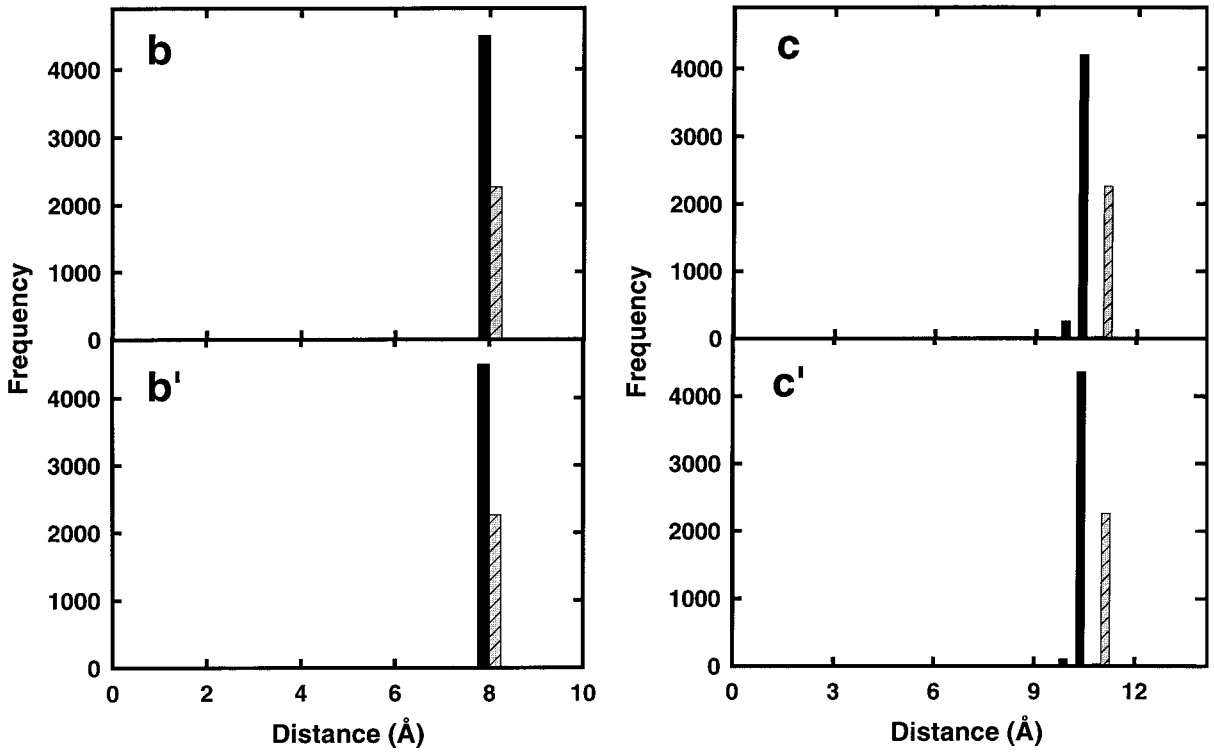


Figure 5. Frequency histograms resulting from the conformational analysis of Red HE-3B. The population of low energy conformers of the dye was described in terms of the parameters shown in Fig. 2b. The distributions of values for parameters b, b′, c and c′ are shown for simulations run in vacuo (dashed grey bars) and in water (black bars).

correlation exists between angles α and β, as revealed by a void occupying a broad central region in a plot (25°–115° and 40°–140°, for α and β, respectively), thus only peripheral values are permitted (results not shown). The significance of these facts lies in that those conformers are favored which exhibit a two-blade propeller shape, where each blade represents the plane of a hemi molecule. Remarkably, the constraint imposed by this cross-correlation disappears in water due to attenuation of repulsive electrostatic interactions. On the other hand, angle γ (or γ′) shows a distribution where most of the values fall in the range 0°–60° (with a maximum at around 30°). This means that line d always orients itself obliquely with respect to the plane of each triangle. Here again, the distribution bell in water is broader than in vacuo.

Angle δ (or δ′) describes the orientation of each triangle in the plane with respect to the projection of line d onto this plane (i.e. the x and x′ axes). By definition, δ (or δ′) ranges between −180° and 180°, i.e. equivalent to a full rotation of segment DA (or D′A′) around D (or D′). Those orientations of DA (or D′A′) pointing approximately to the center of the molecule correspond to negative values of δ (or δ′). Histograms corresponding to results from simulations in vacuo show that most of the conformers exhibit δ (or δ′) values which cluster narrowly around 0° and −140°. This result renders it unlikely that vertex A (or A′) ever points to the center of the molecule (i.e. δ or δ′ ∼ −90°). In the presence of water, the range of these values widen significantly, but these are still centered around the same maxima.

Finally, the internal parameters of each triangle, namely, angle ε (or ε′), and distances b (or b′) and c (or c′) are restricted to the following values: 45°, 7.8 Å and 10.5 Å, respectively. The sharp distributions observed for these parameters validate the initial description of each hemimolecule as a rigid triangle. In addition, these observed values do not differ from those measured in a planar construction (ε = 43.4°, b = 7.8 Å and c = 10.6 Å).

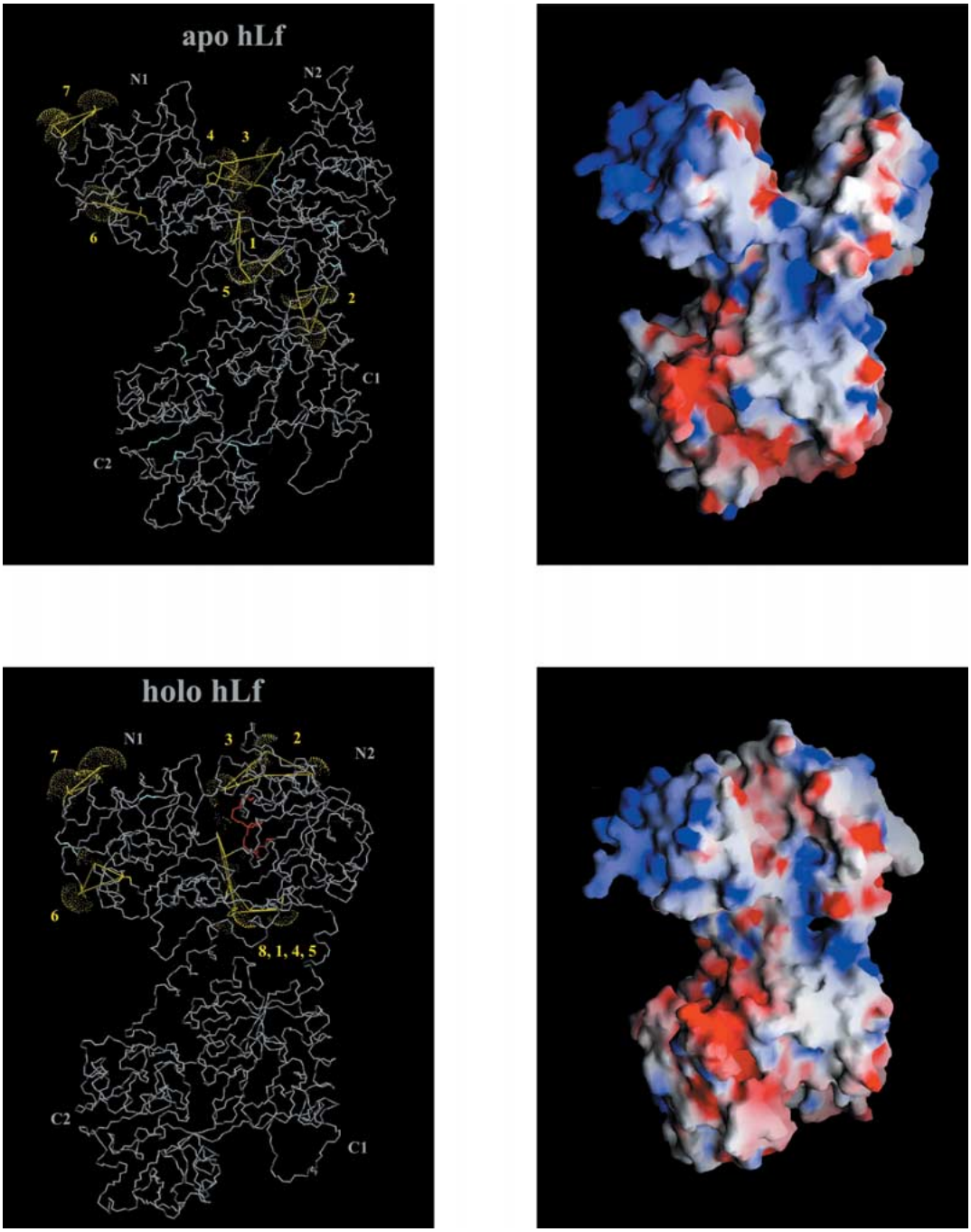


Figure 6. Predicted binding sites for Red HE-3B shown upon the structures of apo and holo hLfs. Left: upon a polytube backbone view of the protein, triangles listed in Tables 3 and 4 are shown in yellow. Dotted caps represent the solvent accessible surface area around positively charged groups. N and C termini of peptides listed in Table 1 are shown in aquamarine. Backbone segments shown in red indicate regions where, due to severe steric hindrance, no binding of the dye would be expected. Right: solvent accessible surface area colored by the electrostatic surface potential of the protein is shown in the same orientation, as calculated by GRASP (fully saturated red and blue correspond to approximately -10 k_BT/q and $+10$ k_BT/q, respectively).

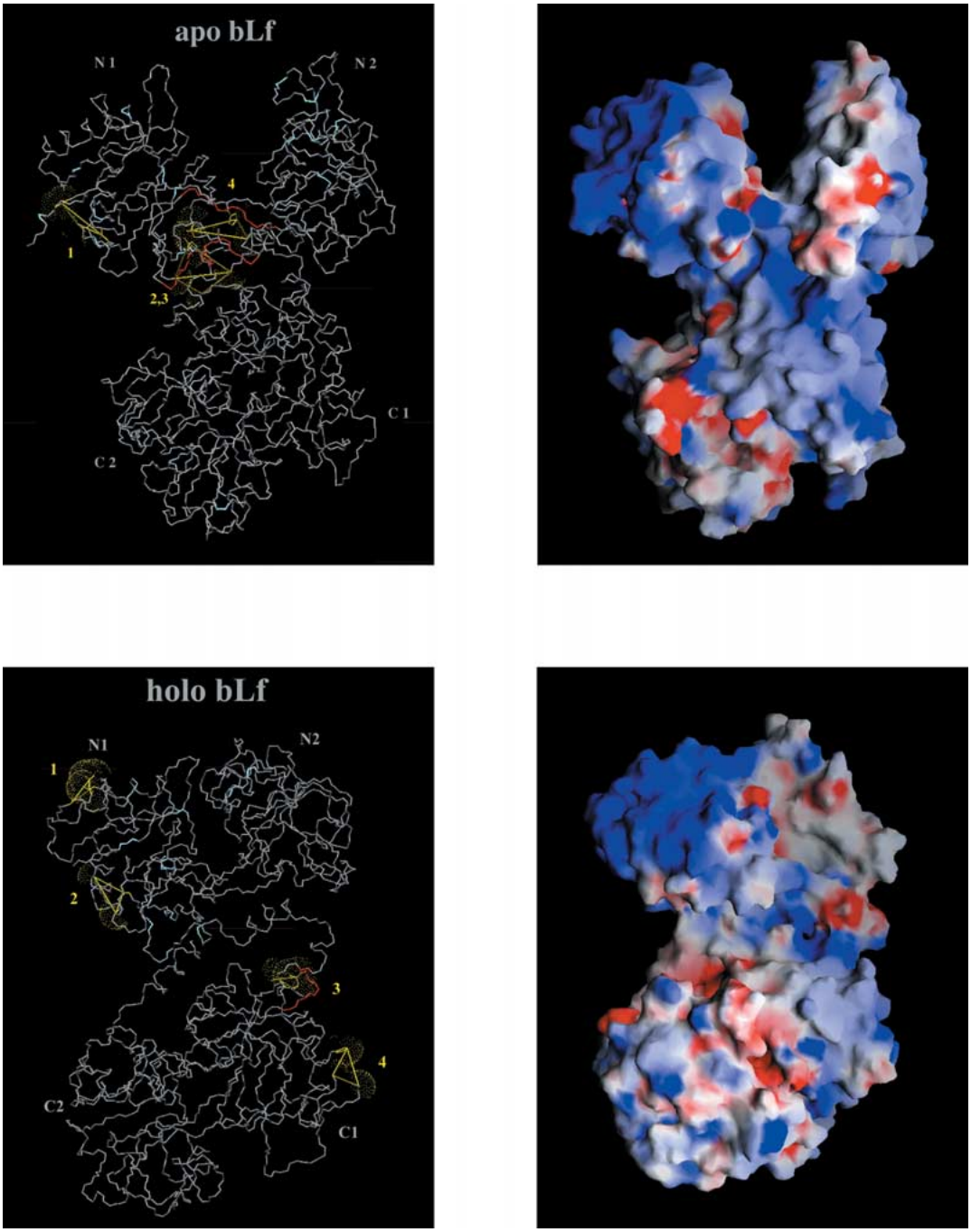


Figure 7. Predicted binding sites for Red HE-3B shown upon the structures of apo and holo bLfs. Left: upon a polytube backbone view of the protein, triangles listed in Tables 5 and 6 are shown in yellow. Dotted caps represent the solvent accessible surface area around positively charged groups. N and C termini of peptides listed in Table 1 are shown in aquamarine. Backbone segments shown in red indicate regions where due to severe steric hindrance, no binding of the dye would be expected. Right: solvent accessible surface area colored by the electrostatic surface potential of the protein is shown in the same orientation, as calculated by GRASP (fully saturated red and blue correspond to approximately -10 k_BT/q and $+10$ k_BT/q, respectively).

Foundation for the interaction model of Red HE-3B and Lf

Mounting evidence points to the fundamental role played by the electrostatic interaction as the main driving force responsible for binding of Red HE-3B to Lf. bLf and hLf – both in their apo- and holoforms- are quantitatively eluted from Red HE-3B immobilized on a Sepharose column if approximately 1.5 M sodium chloride is added [17]. In addition, the analysis of spectral changes (shifts in λ_{max}) in the visible region between free dye and dye-protein complexes allow to determine the prevailing chemical nature of the interaction [26]. According to this criterion, all Red HE-3B/Lf complexes most closely match the case where the electrostatic interaction predominates (Grasselli, M., Cascone O., Anspach F.B. unpubl.).

We envision that the structure of the dye could be considered a chemical scaffold onto which negatively charged groups adopt defined spatial locations. Our contention holds that a favorable face-to-face interaction will occur whenever a complementary pattern of positive charges exists on the protein surface (Fig. 2a).

Given the results described above, Red HE-3B can be regarded as a two-block molecule built from two rigid moieties, each holding in place three negatively charged sulfonate groups, joined by a flexible tether which allows hinge motions and lateral displacements around its center. The ranges of values found for the geometrical parameters impose restrictions on the shape of the molecule and define the most likely conformers present in the ensemble. A set of positively charged groups similarly oriented in space on the protein surface will maximize the favorable electrostatic interaction by minimizing the distances between opposite charges. For all these reasons, the rationale of our approach involves finding every array of positive charges which comply with the geometrical restrictions of the ligand molecule.

In addition, as will becomes evident in some cases, the possibility of steric hindrance occurring at the ligand-protein interface, thus impairing a favorable electrostatic interaction, will also be taken into account by close inspection of each predicted binding site. To achieve our goal we implemented the stepwise procedure outlined below to search for binding sites for the dye on the protein surface.

General strategy for searching binding sites of Red HE-3B on Lf

As stated above, the main criterion for searching binding sites for the dye involves a match between the simplified model of the ligand (considered as a pair of tethered triangles) and a similarly shaped set of opposite charges on the surface of Lfs (Fig. 2a). Every possible set of six positive charges belonging to amino acids with a substantial exposure to the solvent was considered for this analysis. Two subsets of three charges each define surfaces onto which a close fit with the ligand is sought. This is materialized by evaluating the similarity between parameters derived from these surfaces with those calculated from the population of low energy conformers of the dye.

All possible sets of positive charges in the protein potentially interacting with the dye were considered. From the full set of charges, we initially selected those which comply with the following two steps.

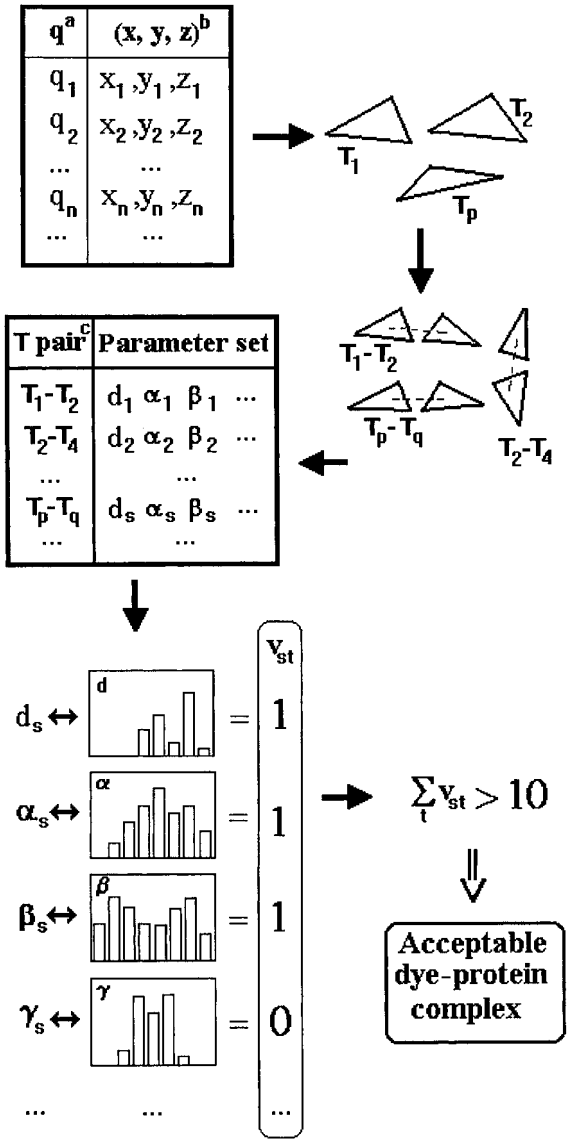
(a) Those parameters which define the shape of each triangle (i.e. b, b′, c, c′, ε and ε′) are the most critical and, therefore, they restrict severely the number of possible binding modes. In addition, to include other plausible models, which would otherwise be discarded, we allowed the relaxation of up to one of these six parameters. The structural implication of this being that not every charge in Red HE-3B should necessary pair with a positively charged amino acid, i.e. the failure to form a single ion pair does not suffice to compromise the binding interaction.

(b) The next restriction involves the choice of a suitable criterion for the selection of pairs of triangles: among all possible pairs, only those which complied with a value of parameter d within a realistic interval were considered (see below).

The general description of the stepwise selection procedure for identifying binding sites is outlined in the scheme and consists of the following steps.

1. Coordinates of Lfs (PDB files: 1LFH, 1LFG and 1BLF, and a model of apo bLf) were used as input for Msroll [32], a program which calculates the solvent accessible surface area of proteins (with a probe radius of 1.5 Å). The output of this analysis identified those basic residues lying at the surface.

2. Cartesian coordinates of the following atom types define the spatial distribution of N positive charges (q_n): NZ for Lys; NE2 for His; and CZ (or NH2 / NH1, depending on which of these atoms is solvent accessible) for Arg.



Scheme 1. Flow diagram of the search strategy for binding sites.

3. Charges (q_n) were clustered into P triangles (T_p) according to the following conditions: (i) 12 Å > (q_j–q_i) > 6 Å and (ii) 12 Å > (q_k–q_i) > 6 Å, where i, j and k varied between 1 and N.

4. A subpopulation of Q triangles was selected according to the following restrictions: (i) 60° > angle q_j–q_i–q_k > 20°; (ii) 9 Å > (q_j–q_i) > 6 Å; (iii) 12 Å > (q_k–q_i) > 8 Å.

5. An initial set of triangle pairs was built using as filter the distance between the centers of mass of the triangles (cm). Only those pairs were chosen where this distance ranges between the following values: 23 Å > (cm(T_p)–cm(T_q)) > 12 Å, where p and q varied between 1 and Q.

6. Cartesian coordinates from each pair of triangles were then used as input for the calculation of the geometrical parameter set defined above (Fig. 2b)

7. A Boolean variable (v_{st}) was defined for each parameter t belonging to the pair of triangles s. This variable equals one whenever a given geometrical parameter falls within the allowed interval stated below, otherwise it was set to zero. The limits for each interval were established on the basis of the distribution of values calculated from the conformational analysis of the free dye (see Figs. 3, 4 and 5): (i) 23 Å > d_s > 12 Å; (ii) 130° > α_s > 20°; (iii) 150° > β_s > 40°; (iv) 60° > γ_s > 10°; (v) 60° > $\gamma_{s'}$ > 10°; (vi) −130° > δ_s > −180° or 60° > δ_s > −60° or 180° > δ_s > 130°; (vii) −130° > $\delta_{s'}$ > −180° or 60° > $\delta_{s'}$ > −60° or 180° > $\delta_{s'}$ > 130°; (viii) 50° > ε_s > 30°; (ix) 50° > $\varepsilon_{s'}$ > 30°; (x) 9 Å > b_s > 6 Å; (xi) 9 Å > $b_{s'}$ > 6 Å; (xii) 12 Å > c_s > 9 Å; (xiii) 12 Å > $c_{s'}$ > 9 Å.

8. A pair of triangles s was selected if it complies with the condition that the sum of v_{st} on t (defined as the quality score Σv) be 13. In addition, those pairs of triangles were also selected which fail to comply with the allowed values in no more than two parameters (i.e. where $\Sigma v = 12$ or $\Sigma v = 11$), but where one and only one of these parameters belong to those defining the triangles (i.e. b, b′, c, c′, ε or ε′). Finally, this combined subset of triangle pairs describes the plausible binding sites for the dye. The results from this analysis are summarized in Tables 2–6 and Fig. 6 and 7.

Constraints of the search strategy

The input population of positively charged groups in Lfs includes all those belonging to exposed basic residues, regardless of the extent of their exposure to the solvent. At this stage, no attempt was made to filter off groups involved in any possible ion pairs.

The strategy described above analyzes the spatial correspondence of a set of six positively charged groups on the surface of the protein with the geometry of the most likely conformers of the dye. From the start, this assumption excludes putative sites where less than six charges participate. However, one should be reminded that, in the end, the procedure includes a relaxation of this constraint by allowing a mismatch of at most two parameters (see point 8 in the previous section). The contention here is that, although less than six charges would permit binding to occur,

a full match might be required to achieve a proper interaction.

On the other hand, up to this point this method ignores possible steric clashes between protein atoms and the central linker tethering both moieties of Red HE-3B. For this reason, all putative binding sites resulting from the output of this search (Tables 3–6) were further examined after graphical rendering of triangle pairs upon the protein structure (Fig. 6 and 7). Situations where a significant volume overlap (steric hindrance) is expected between the ligand and protein were ruled out.

Protection from proteolysis in Red HE-3B/Lf complexes

The experimental counterpart of this study involved looking for binding sites of the dye by protection from enzymatic cleavage on Red HE-3B/Lf complexes. Currently a wealth of information exists on structural data of ligand-protein complexes: 1019 complexes were reported in PDB (www.rcsb.org) up to the year 2000 according to Moreno and Leon [33]. However, scant knowledge is available on the nature of the interaction between triazine dyes and binding sites in proteins. Most of the known complexes involve cases where these dyes interact due to their structural similarity to coenzymes (e.g. NADP). In this regard, so far no coenzyme binding sites were described in Lfs. Despite this fact, Lfs are known to bind tightly triazinic dyes ($K_d < 1$ μM for Red HE-3B: Grasselli, M., Cascone O., Anspach F.B. unpubl.) to multiple sites. This compounds the difficulty of obtaining suitable crystalline complexes for analysis. In addition, as has been discussed above, the prevailing electrostatic nature of the dye-protein interaction precludes the use of high salt concentration which is common in crystallization. Thus, protection from proteolysis appears as a promising alternative method from which to derive structural data. To this end, the relatively large size of Red HE-3B (a molecule 25 Å long and ∼7.3 Å wide) warrants the use of this technique, which – at first sight – appears more appropriate to analyze protein-protein or protein-peptide interfaces. In fact, the measured solvent accessible surface area of Red HE-3B (∼1,473 Å^2) compares favorably with 'standard-size' interfaces (1,600 ± 400 Å^2), such as those found in protein-inhibitor complexes [34].

On the other hand, although we cannot completely rule out the existence of conformational changes induced by the dye-ligand interaction, these changes – if any – would be expected to be only small [34]. Thus, the data presented below is interpreted as resulting directly from the primary binding event occluding cleavage sites to the attack by proteases. The facts that (i) Red HE-3B is not – and does not bear any similarity to – a natural ligand of Lfs, and (ii) the extreme flexibility of this molecule which would allow it to match 'softly' to differently shaped areas would support the above assumption. In this regard, it has been shown that Lfs – both in their apo- and holo- forms – are recovered in high yields and in their functional forms after affinity chromatography on immobilized Red HE-3B columns.

Tryptic digests of both the isolated proteins versus the dye-protein complexes (bLf and hLf in both their apo and holo forms) were compared after separation of peptides by reversed-phase HPLC. Conditions for digestion were adjusted so as to match as closely as possible those of the binding assay (10 mM sodium phosphate buffer, pH 7.25), thus allowing a direct comparison between binding and proteolysis protection data. Nevertheless, the conditions chosen were more severe than those employed to excise domains and which served to differentiate apo from holo forms [35]. In our case, apo as well as holo-Lfs were digested to yield peptides of less than 10 kDa. Similarly to previous work [35], holo-hLf proved to be most resistant to hydrolysis.

The HPLC separation pattern of the tryptic digest of each Lf revealed about 30 major peaks, whereas in those of each Red HE-3B/Lf complex, at least 5 peaks were absent, thus indicating that the dye indeed protected some peptide bonds from tryptic cleavage. In this situation, the interaction site could be located at any or both termini of the peptide. After microsequencing analysis we were able to identify those peptides listed in Table 1. In general, the holo forms of the proteins yielded fewer peptides than the apo forms, a feature possibly related to the more compact structure exhibited by the former (see below Fig. 6). For bLf, 6 and 8 peptides were identified for the holo and apo forms, respectively. Among these peptides, 4 were common. By contrast, hLf revealed fewer differences, although this fact might have arisen from the intrinsically lower yield of enzymatic hydrolysis compounded by difficulties associated with ambiguity in the interpretation of sequence data.

Locations predicted for binding sites of the dye and their correlation with proteolysis-protected regions

Table 2 summarizes the progress of the search procedure for binding sites for Red HE-3B in all Lfs assayed. In turn, Tables 3–6 present the output of this computational analysis. Each predicted binding site, listed in the first column in Tables 3–6, is described by a pair of triangles. Rows list: (i) the values of the internal parameter set which defines the geometry of the binding site, and (ii) the quality score Σv.

The computational procedure predicted only a limited number of binding sites: 2–7 such sites were found out of a total of 97–98 basic amino acids (Table 2). Fig. 6 and 7 show: (i) polytube diagrams of the 4 proteins dealt with in this study, where experimentally determined sites protected from enzymatic cleavage (in aquamarine) and the predicted interaction areas with Red HE-3B (yellow triangles showing the solvent accessible surface area of positively charged groups) are represented; and (ii) electrostatic potential maps at the protein surface as rendered by GRASP [36]. Numbered triangles correspond to those defined in Tables 3–6.

The proximity between predicted and actual binding (protected) sites was evaluated by measuring the distance between any of the two centers of mass of each triangle and the Cα of the corresponding amino acid.

Apo hLf

Fig. 6 shows three possible binding modes for the dye in apo hLf, two of which partially overlap (see Table 3). Two of the predicted binding sites map to the region connecting subdomains N1 and N2, which in this form are further apart than in holo hLf. In fact, triangles 3 and 4 are close to amino acids which constitute the iron-carbonate binding site in this domain of the holo form. Triangle 2 bridges a cleft running between two main lobes, whereas triangle 1 is located at the junction between N1 and N2 sublobes adopting an orientation roughly perpendicular to the protein surface. In spite of this drawback, triangle 5, which lies nearby, has the ability to form a pair with triangle 2. No evident steric hindrance occurs with other residues preventing the fitting of the linker region of Red HE-3B between triangle pairs 1–2, 3–5, 4–5 and 2–5. Accordingly, this gorge-like region half-encircling sublobe N2 shows defined positive character and ready accessibility for dye binding.

On the other hand, triangles 6 and 7 are located on opposite sides of a very convex and densely positive surface in sublobe N1 where distance d is extreme (22.9 Å). Therefore it becomes difficult to bridge them in any reasonable fashion, thus rendering it unlikely as a putative binding site. Nevertheless, higher mobility at the N terminal end, where R2, R3 and R4 are located could possibly give rise to a new binding site in this region.

In summary, two partially overlapping binding regions would result from this analysis, namely one involving triangles 1–2 (or 2–5) and a second one involving triangles 3–5 (or 4–5). In agreement with this prediction, three sites in subdomain N2 protected from proteolysis were found in close proximity (at less than 11 Å from any of the vertices, see footnote[a]) to these putative binding regions.

Holo hLf

In this case the search produces as output seven tentative binding modes for the dye (Fig. 6 and Table 4). Similarly to apo hLf, these predicted binding sites are located at the N terminal domain and can be clustered into two regions which map to the N1 subdomain (triangles 6–7) and to the N2 subdomain or the contact zone (the remaining triangles). However, inspection of the latter shows that the amino acid stretch I209-L218, comprising a short beta strand and alpha helix 7, would pose severe steric hindrance in between triangles sets 2, 3 and 1, 4, 5, 8. Besides, some of the triangles in the second group occur in a narrow cleft which will certainly prevent the access to the ligand. In conclusion, triangle pair 6–7 is left as the only putative binding site for Red HE-3B. In this N1 subdomain, cleavage sites 31 and 38 (Table 1) lie at <15 Å and <20 Å, respectively. This site differs from that found in the apo form in that here R3, instead of R2, would be involved in the interaction. Although the N1 sublobes of each form are almost superimposable, the N terminus end (i.e. peptide G1-S5) adopts a different orientation, thus explaining the result above.

Apo bLf

Since no structure is available for apo bLf, a model was constructed upon the cristallographic coordinates of apo hLf taking advantage of the high degree of similarity between these proteins (70% identity in the amino acid sequence). Except for a minor difference

[a]A distance of 11 Å is shorter than the minor radius of an ellipsoid representing the smallest subdomain (or lobe) of Lf.

Table 2. Statistics findings of the search algorithm for binding sites[a]

Set size	apo hLf	holo hLf	apo bLf	holo bLf
RKH[b]	97	97	98	98
Solvent-exposed RKH[c]	88	86	93	91
Triangles[d]	59	57	44	55
Selected triangles[e]	20	22	6	16
Triangle pairs[f]	32	23	3	14
High-scoring triangle pairs[g]	5	7	2	2

[a]A detailed description of the procedure is presented in section *General strategy for searching binding sites of Red HE-3B on Lf*.
[b]Basic amino acids in Lfs.
[c]Subset of basic amino acids as defined in step 2.
[d]As defined in step 3.
[e]As defined in step 4.
[f] As defined in step 5.
[g] As defined in step 8.

Table 3. Geometrical description of putative binding sites for Red HE-3B in apo hLf[a].

Triangle pair[b]	d (Å)	α (°)	β (°)	γ (°)	γ′ (°)	δ (°)	δ′ (°)	ε (°)	ε′ (°)	b (Å)	b′ (Å)	c (Å)	c′ (Å)	Σv
1–2	18.1	89	45	43	11	49	39	29	38	6.4	8.2	9.9	9.8	12
3–5	18.8	89	53	37	32	34	5	52	41	8.6	6.8	10.4	9.2	12
4–5	18.0	92	129	20	44	144	9	23	41	8.6	6.7	10.3	9.2	12
2–5	13.6	52	94	31	81	37	86	38	41	8.2	6.8	9.8	9.2	11
6–7	22.9	59	39	32	34	10	165	38	53	8.7	7.9	9.3	11.2	11

[a]Values are listed for the geometrical parameters and quality score as defined in the section *General strategy for searching binding sites of Red HE-3B on Lf*.
[b]Triangles are defined by the coordinates of the atoms defined in step 2 of section *General strategy for searching binding sites of Red HE-3B on Lf*. Triangle numbers (and the aminoacids defining them) are as follows: 1 (R89,H91,H243); 2 (K661,K237,K241); 3 (R121,R210,K301); 4 (R210,H253,K301); 5 (K243,H246,R690); 6 (R2,R258,K263); 7 (R24,R27,K28). This numbering convention corresponds to that shown in Fig. 6.

Table 4. Geometrical description of putative binding sites for Red HE-3B in holo hLf[a].

Triangle pair[b]	d (Å)	α (°)	β (°)	γ (°)	γ′ (°)	δ (°)	δ′ (°)	ε (°)	ε′ (°)	b (Å)	b′ (Å)	c (Å)	c′ (Å)	Σv
1–2	22.8	84	121	37	42	−54	14	40	39	6.5	9.0	10.3	10.1	13
1–3	20.2	105	122	24	38	−57	−120	40	34	6.5	8.0	10.3	11.3	12
4–3	22.9	103	60	21	43	−45	−132	44	44	6.7	8.0	11.7	11.3	13
5–3	22.9	60	138	27	41	−133	−137	44	44	7.0	8.0	11.7	11.3	13
6–7	22.0	104	92	27	49	−175	−85	43	56	6.4	7.9	9.4	9.3	11
8–2	19.1	86	114	45	39	138	26	30	39	6.5	9.0	9.5	10.2	13
8–3	15.9	107	117	30	34	134	97	30	44	6.6	6.5	9.5	11.3	12

[a]Values are listed for the geometrical parameters and quality score as defined in the section *General strategy for searching binding sites of Red HE-3B on Lf*.
[b]Triangles are defined by the coordinates of the atoms defined in step 2 of section *General strategy for searching binding sites of Red HE-3B on Lf*. Triangle numbers (and the aminoacids defining them) are as follows: 1 (R89,H91,R210); 2 (K99,K197,K200); 3 (K197,K200,K296); 4 (R89,H91,H246); 5 (R89,K243,H246); 6 (R3,R258,K263); 7 (R24,R27,K28); 8 (R89,R210,K301). This numbering convention corresponds to that shown in Fig. 6.

Table 5. Geometrical description of putative binding sites for Red HE-3B in model of apo bLf[a].

Triangle pair[b]	d (Å)	α (°)	β (°)	γ (°)	γ′ (°)	δ (°)	δ′ (°)	ε (°)	ε′ (°)	b (Å)	b′ (Å)	c (Å)	c′ (Å)	Σv
1–2	22.9	124	112	12	39	133	18	47	57	6.6	8.5	9.0	11.6	12
3–4	23	22	107	64	46	1	173	32	59	8.9	8.2	9.9	9.9	11

[a]Values are listed for the geometrical parameters and quality score as defined in the section *General strategy for searching binding sites of Red HE-3B on Lf.*
[b]Triangles are defined by the coordinates of the atoms defined in step 2 of section *General strategy for searching binding sites of Red HE-3B on Lf.* Triangle numbers (and the aminoacids defining them) are as follows: 1 (R12,K263,R268); 2 (R74,K78,K334); 3 (H96,K248,R694); 4 (K78,K334,K391). This numbering convention corresponds to that shown in Fig. 7.

Table 6. Geometrical description of putative binding sites for Red HE-3B in holo bLf[a].

Triangle pair[b]	d (Å)	α (°)	β (°)	γ (°)	γ′ (°)	δ (°)	δ′ (°)	ε (°)	ε′ (°)	b (Å)	b′ (Å)	c (Å)	c′ (Å)	Σv
1–2	22.3	65	46	46	14	−137	−68	52	38	6.2	6.3	11.5	9.7	11
3–4	19.9	52	119	44	0	−19	6	31	51	6.1	7.2	9.3	11.9	11

[a]Values are listed for the geometrical parameters and quality score as defined in the section *General strategy for searching binding sites of Red HE-3B on Lf.*
[b]Triangles are defined by the coordinates of the atoms defined in step 2 of section *General strategy for searching binding sites of Red HE-3B on Lf.* Triangle numbers (and the aminoacids defining them) are as follows: 1 (R20,R25,K27); 2 (K52,R258,K263); 3 (R341,R344,R603); 4 (H613,K623,K628). This numbering convention corresponds to that shown in Fig. 7.

in a short loop region (KSSKHS in this model, instead of SQQSSD in apo bLf) the backbones of these two structures are essentially super imposable.

Figure 7 and Table 5 show the model of apo bLf where two partially overlapping sites are described. Triangles 3 and 4 occur in the N domain on opposite sides along the central cleft, therefore rendering them impossible to connect in any reasonable way (a short α helix K318-S327 and an extended chain between A253 and H258 stand in between them). On the other hand, triangle pair 1–2 emerges as a plausible binding site. This site belongs to the N1 sublobe. Here, the N-terminal segment extending up to R8 (in particular residue P7) partially covers the outer edge 12–268 in triangle 1. However, the increased mobility expected for this region will likely alleviate any consequent steric hindrance. Moreover, local uncertainty exists in the model (the initial pentapeptide was not modeled at all). By contrast, triangle 2 is fully solvent exposed. Accordingly, most sites protected from enzymatic cleavage cluster in the N1 subdomain near the predicted binding site.

Holo bLf

In holo bLf (Fig. 7 and Table 6) only two non-overlapping sites are described. The prediction for holo bLf holds that two different regions of the protein surface could be involved in ligand binding, namely those mapped by triangle pairs 1–2 and 3–4 in the N1 and C1 sublobes, respectively. Nevertheless, the C terminal region of helix 13 connecting the N and C domains (peptide 341–345) would interpose itself between triangles 3 and 4, thus rendering this site unlikely for binding. Similarly to holo hLf proteolysis protected sites cluster mostly in the N1 subdomain (W8, K18, R47, L74 are located at less than 15 Å from this site, and I310 and V314 at less than 20 Å), close to the predicted binding site 1–2. In addition, W24 occurs near the center of triangle 1, orienting its indole ring approximately parallel to its surface, thus possibly contributing a favorable stacking interaction with the aromatic dye.

In general, in its current version the method predicts a limited number of sites, i.e. more proteolysis protected sites are experimentally found than anticipated by predictions. From the experimental standpoint, a comparison of HPLC patterns does not warrant having found all possible peptides involved in binding regions, thus causing a consequent lack of completeness of the data set. On the other hand, an intrinsically different susceptibility to proteolysis of each protein species indeed exists, e.g. apo hLf

was found to be significantly more resistant to tryptic proteolysis [35]. In addition, certain sites might appear in a time-dependent fashion as a consequence of the previous cleavage at others, thus blurring a straightforward correlation to be made, e.g. a differential tendency to proteolysis at the N terminal lobe could determine the appearance of cleavage sites at the C domain, as becomes readily apparent in the apo forms. This is particularly evident in that, although an excellent superimposition exists between the slightly negative C terminal domains of apo and holo hLf, the appearance of protected sites varies widely.

Independent evidence from differential visible absorption measurements run under the same conditions of pH and ionic strength as those used for proteolysis indicates that between three and four interaction sites for the dye were found in hLf and bLf (regardless of their form), respectively (Grasselli, M., Cascone O., Anspach F.B. unpubl.).

Conclusions and perspectives

As regards the algorithm presented herein, any quantitative correlation with experiment aimed at evaluating the efficiency of the search should take into account the intrinsic tunable stringency of the method. This allows a limited number of geometrical violations to take place on demanding a particular spatial orientation to the array of charges involved, meaning that other (more distorted) sites could appear if increased relaxation is allowed. In order to maximize the predictive power of this method without compromising the accuracy of the results, in each particular case one could in principle introduce modifications such as including weighting factors to establish a hierarchy among the set of parameters, or selecting those which best represent the geometry of the molecule. In our case, the applied strategy involved the relaxation of a single triangle-related parameter which produced a reasonable agreement with experimental data while avoiding the appearance of false positive results.

In principle, one could envision that this method could bear relevance for the screening of possible dye-binding sites in proteins whose structure is known whenever the electrostatic interaction prevails. Results derived thereof could serve as a suitable starting point for further refinement by classical docking analysis. In addition, its use as a predictive tool for binding interactions could assist the purification by dye-affinity chromatography of engineered protein variants.

Materials and methods

HLf , bLf and nitrilotriacetic acid were obtained from Fluka (Neu-Ulm, Germany). 2-Mercaptoethanol and trypsin were obtained from Sigma (Deisenhofen, Germany); dithiothreitol from E. Merck (Darmstadt, Germany); and Sephadex LH-20, Heparin-Sepharose CL-6B, and PD-10 columns from Pharmacia (Freiburg, Germany). Vilmax Red HE-3B (Fig. 1) was a generous gift from Dr. J. Mazza, Vilmax (Buenos Aires, Argentina).

Molecular modeling

Model building and energy calculations of Red HE-3B were carried out with MacroModel 5.5 and BatchMin 5.5 [29] installed on a Silicon Graphics O2 workstation (R10000, 320 MB RAM, 54 GB hard disk) under the Irix 6.3 operating system. We used the MM2* force field, the version of Allinger's MM2 force field as implemented in MacroModel. By default, atomic partial charges were calculated from data in the molecular mechanics force field chosen. The MM2* force field uses distance-dependent dielectric electrostatics, instead of the standard dipole-dipole electrostatics. The conformational search of Red HE-3B was carried out following an optimized Monte Carlo method [30]. The electrostatic cut-off was set to 50 Å. For energy minimization, we used a conjugate gradient method, with a final gradient value of 0.05 kJ $Å^{-1}$ mol^{-1} (0.01 kcal $Å^{-1}$ mol^{-1}) as the criterion for convergence. After 50000 Monte Carlo steps, the resulting set of conformers were fully minimized and compared. In the end, all non-identical and non-enantiomeric conformers within a 25 kJ/mol (6 kcal/mol) window above the global minimum were tabulated.

A molecular model for apo bLf was built by homology modeling upon the known structure of apo hLf (PDB file 1LFH) using the automatic mode implemented in SwissModel [37]. Surface residues were identified with the MSP software [32].

The reparametrization of the coordinates corresponding to the structures of the dye and the subsequent search algorithm for binding sites on protein surfaces were implemented with macros built on Microsoft Excel. In its present implementation, our search method is fully automated. The initial input worksheet consists of the set of coordinates of positively charged atoms located on the surface of the protein, as defined in step 2 of the section *General strategy for searching binding sites of Red HE-3B on*

Lf. A linear sequence of macros is then executed, where each one processes the output worksheet produced by the previous one. Tables 3–6 report the output of the final worksheet.

All molecular images were rendered in polytube representation with MacroModel (Mohamadi et al. 1990) and the accessible surface areas showing electrostatic potentials were drawn with GRASP [36].

Purification of Red HE-3B

The crude dye was dissolved in methanol:water (50:50, v/v) and deactivated with 2-mercaptoethanol. Insoluble material was removed by passage through a 0.45 μm filter. The dye was purified by chromatography on a Sephadex LH-20 column equilibrated with the same solvent [38]. The eluate was analysed by TLC on activated silica gel developed with 1-butanol: 2-propanol: ethyl acetate: water (2:4:1:3, v/v).

The concentration of Red HE-3B was determined spectrophotometrically based on the reported molar absorption coefficient (30,000 L mol^{-1} cm^{-1} at 530 nm in aqueous solution at pH 7.0) [25].

Preparation of apo forms of Lf

Lactoferrin (0.5 ml of bLf or hLf, 2 mg/ml dissolved in 10 mM citric acid) was dialysed against two changes of this buffer for 24 h. Finally, protein solutions were exhaustively dialysed against water and lyophilized.

Preparation of holo forms of Lf

The holo form of hLf was obtained as described before [39]. Briefly, hLf (5 mg) was dissolved in 10 mM Tris/HCl, pH 7.3 (250 μl). Freshly prepared 10 mM $NaHCO_3$ was then added (100 μl) and, subsequently, 10 mM Fe(III)-NTA complex was mixed with constant stirring (4 × 50 μl). After 30 min, the reaction mixture was loaded on a Heparin-Sepharose CL-6B column (30 × 5 mm), equilibrated with 10 mM sodium phosphate buffer, pH 7.25. The hLf fraction eluted after addition of 1 M NaCl and was then desalted against 10 mM sodium phosphate buffer by using a PD-10 column.

The holo form of bLf was obtained as described before [40]. Briefly, bLf (5 mg) was dissolved in 50 mM $NaHCO_3$, pH 7.5 (450 μl) and a 10 mM Fe(III) ammonium citrate solution was then added (10 μl). After incubation at 37 °C overnight, the solution was applied to the Heparin-Sepharose CL-6B column as described above for hLf.

Iron saturation of lactoferrin in these samples was ascertained by measuring absorbance ratios at defined wavelengths, namely, the value of A_{280}/A_{466} should lie between 20–22 and that of A_{412}/A_{466} between 0.70–0.74 for hLf [41]. Correspondingly, the ratio A_{280}/A_{465} should be less than 27 and A_{410}/A_{465} should lie between 0.80 and 0.85 for bLf [42].

Preparation of dye-protein complexes

Lactoferrin (0.5 ml of bLf or hLf, 2 mg/ml dissolved in 20 mM sodium phosphate buffer, pH 7.25) was incubated with 0.5 mM Red HE-3B (0.5 ml). Free dye was removed by passage through a PD-10 column and the dye–protein complex was concentrated by ultrafiltration on a Centricon 10 concentrator (10,000 MW cut-off, Amicon, Danvers, MA).

Tryptic digestion of the dye-protein complexes

Each Red HE-3B/Lf complex (1 mg) was incubated at 37 °C in 10 mM sodium phosphate buffer pH 7.3 with trypsin (1:40 enzyme/substrate molar ratio). After 3 h, an equal amount of enzyme was added, and the incubation was continued for 12 h. The resulting low molecular weight peptides (<10 kDa) were separated from the reaction mixture by ultrafiltration through a Centricon 10 concentrator. Control proteolysis experiments using Lfs were run under identical conditions.

Peptide patterns were analysed after separation of the ultrafiltrates by reversed-phase HPLC (Vydac 218TP54 column, 4.6 × 150 mm) at 40 °C using a Waters liquid chromatograph. A linear gradient (0.06% aqueous TFA / 80% acetonitrile in 0.05% aqueous TFA, 90:10 to 20:80, v/v) was developed for 80 min at a flow rate of 0.5 ml/min. Peptide elution was monitored by measuring the absorbance at 215 nm.

Amino acid sequence analysis

Identification of selected peptide peaks was achieved after automated Edman degradation on an Applied Biosystems Model 473A gas phase protein sequencer.

Acknowledgements

M.G., O.C. and J.M.D. thank support from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad de Buenos Aires (UBACyT) and Agencia Nacional de Promoción Científica y Tecnológica (ANPCYT).

The spreadsheet and macros used in the search of binding sites are available from M.G. upon request.

References

1. Baggiolini, M., De Duve, C., Masson, P.L. and Heremans, J.F., J. Exp. Med., 131 (1970) 559.
2. Masson, P.L., Heremans, J.F. and Dive, Ch., Clin. Chim. Acta, 14 (1966) 735.
3. Reiter, B., Int. J. Tissue React., 5 (1983) 87.
4. Ellison, R.T. and Giehl, T.J., J. Clin. Invest., 88 (1991) 1080.
5. Yamauchi, K., Tomita, M., Giehl, T.J. and Ellison, R.T., Infection and Immunity, 61 (1993) 719.
6. Baker, E.N., Rumball, S.V. and Anderson, B.F., TIBS, 12: (1987) 350.
7. Anderson, B.F., Baker, H.M., Norris, G.E., Rice, D.W. and Baker, E.N., J. Mol. Biol., 209 (1989) 711.
8. Norris, G.E., Anderson, B.F. and Baker, E.N., Acta Crystallogr. Sect. B, 47 (1991) 998.
9. Moore, S.A., Anderson, B.F., Groom, C.R., Haridas, M. and Baker, E.N., J. Mol. Biol., 274 (1997) 222.
10. Gerstein, M., Anderson, B.F., Norris, G.E., Baker, E.N., Lesk, A.M. and Chothia C., J. Mol. Biol., 234 (1993) 357.
11. Pahud, J.J. and Hilpert, H., Prot. Biol. Fluids, 23: (1976) 571.
12. Kawakami, H., Shinmoto, H., Dosako, S.I. and Sogo, Y., J. Dairy Sci., 70 (1987) 752.
13. Yoshida, S., J. Dairy Sci., 72 (1989) 1446.
14. Yoshida, S. and Xiuyun, Y., J. Dairy Sci., 74 (1991) 1439.
15. Yoshida, S. and Xiuyun, Y., Neth. Milk Dairy J., 45 (1991) 273.
16. Bezwoda, W.R. and Mansoor, N., Clin. Chim. Acta, 157 (1986) 89.
17. Grasselli, M. and Cascone, O., Neth. Milk Dairy J., 50 (1996) 551.
18. Scopes, R.K., Analyt. Biochem., 165 (1987) 235.
19. Vijayalakshmi, M.A. and Bertrand More, O. (Eds.) Protein-dye interactions: Developments and Applications, Elsevier Applied Science, London, 1989, p. 265.
20. Boyer, P.M. and Hsu, J.T., Adv. Biochem. Eng. Biotechnol., 49 (1993) 1.
21. Biellmann, J.F., Samama, J.P., Branden, C.I. and Eklund, H., Eur. J. Biochem., 102 (1979) 107.
22. Dean, P.D.G. and Watson, D.H., J. Chromatogr., 165 (1979) 301.
23. Shimazaki, K. and Nishio, N., J. Dairy Sci., 74 (1991) 404.
24. Shimazaki, K., Nitta, K., Sato, T., Tomimura, T. and Tomita, T., Comp. Biochem. Physiol., 101B (1992) 541.
25. Cadelis, F., Kirchberger, J., Vijayalakshmi, M.A. and Kopperschlager, G., Biomed. Biochim. Acta, 50 (1991) 1167.
26. Subramanian, S., Arch. Biochem. Biophys., 216 (1982) 116.
27. Barden, R.E., Darke, P.L., Deems, R.A. and Dennis, E.A., Biochemistry, 19 (1980) 1621.
28. Halperin, I., Ma, B., Wolfson, H. and Nussinov, R., Proteins, 47 (2002) 409.
29. Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T. and Still, W.C., J. Comput. Chem., 11 (1990) 440.
30. Chang, G., Guida, W.C. and Still, W.C., J. Am. Chem. Soc., 111 (1989) 4379.
31. Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., J. Am. Chem. Soc., 112 (1990) 6127.
32. Connolly, M.L., J. Mol. Graphics, 11 (1993) 139.
33. Moreno, E. and Leon, K., Proteins, 47 (2002) 1.
34. LoConte, L., Chothia, C. and Janin, J., J. Mol. Biol., 285 (1999) 2177.
35. Brines, R.D. and Brock, J.H., Biochim. Biophys. Acta, 759 (1983) 229.
36. Nicholls, A., Sharp, K. and Honig, B., Proteins Structure, Function and Genetics, 11 (1991) 281.
37. Guex, N. and Peitsch, M.C., Electrophoresis, 18 (1997) 2714.
38. Easton, M.J. and Yon, R.J., Biochim. Biophys. Acta, 1118 (1992) 298.
39. Graham, G. and Bates, G.W., J. Lab. Clin. Med., 88 (1976) 477.
40. Shimazaki, K. and Hosokawa, T. A., Anim. Sci. Technol. (Jpn.), 62 (1990) 354.
41. Norris, G.E., Baker, H.M. and Baker, E.N., J.Mol.Biol., 209 (1989) 329.
42. Brown, E.M. and Parry, R.M., Biochemistry, 13 (1974) 4560.