Biophysical Chemistry 9 (1979) 383–391

MULTISITE PROTEINS AND RANDOMLY COILED POLYMERIC LIGANDS. CHAIN LENGTH DEPENDENCE OF BINDING CONSTANTS☆

Stefan STANKOWSKI and JürgenENGEL
Department of Biophysical Chemistry, Biozentrum der Universitat Basel, CH-4056 Basel, Switzerland

Richard A. BERG *
Department of Biochemistry, College of Medicine and Dentistry of New Jersey, Rutgers Medical School, Piscataway, NJ 08854, USA

Received 6 January 1979

A model mechanism was developed for the binding of a rigid multisite protein with a randomly coiled multivalent ligand. Probabilities of the formation of chain loops between sites located at given distances at the protein were calculated by an extension of the concept of ring closure in coiled chain molecules. Expressions were derived for the dependence of overall equilibrium quantities, such as the binding constant between the protein and the ligand, on intrinsic parameters such as intrinsic binding constants, number of sites at the protein and their distances and on the chain length of the polymeric ligand. A pronounced chain length dependence of the overall binding constant was predicted even at chain lengths much longer than the size of the protein. Such a dependence was previously observed for the enzyme prolyl hydroxylase which acts on polymeric substrates like $(ProProGly)_n$. This so far unexplained feature is quantitatively described by the model mechanism which is believed to be applicable to many other interactions of biological importance.

1. Introduction

An essential feature of enzymes which modify polymeric substrates is that they often must react repeatedly with the same substrate molecule. The interaction may be sequential as in the case of nucleic acid polymerases [1–6], RNase [7] and glycogen synthetase [8] where the product of one catalytic event serves as the substrate for a succeeding catalytic event. Alternatively multiple interactions between an enzyme and a polymeric substrate may occur as independent events where different sites or segments of the substrate are modified independently. Examples of enzymes operating in such a manner are nucleic acid methylases [9], glycosyl transferases [10–12], prolyl hydroxylase [13,14] and lysyl hydroxylase [14,15]. Multiple interactions with a flexible polymer chain may occur for proteins other than enzymes such as the binding of IgG to polysaccharides [16,17].

In this report a mechanism is presented which may be of general importance for the binding of a multivalent molecule to flexible polymeric ligands. The mechanism is based on the concept of ring closure which was developed to describe cyclization reactions and the interaction of chain segments in a polymer to which reactive and catalytic groups are bound (for review see ref. [18]). A mathematical treatment has previously been formulated to estimate the probability of ring closure for two segments of a polymer [19,20]. In the mechanism developed here the concept of ring closure has been extended to include multiple loops which are not completely closed because of the distance between binding sites on the multivalent protein. The mechanism results in an enhanced binding affinity between a multivalent protein (enzyme) and a flexible ploymeric ligand (substrate). The binding affinity is crucially dependent on the length and flex-

☆ This work was supported in part by research grant AM-16,526 from the National Institute of Health.
* Present address: Pulmonary Branch, National Heart, Lung, and Blood Institute, Bethesda, Maryland 20014, USA.

ibility of the polymer chain. Such a mechanism can be used to describe the unusual chain length dependence of Michaelis constants for substrates of prolyl hydroxylase [14,21–24].

2. Probability of loop formation

2.1. Selective binding

Let us first consider the case of a multivalent protein (enzyme E) which interacts with a polymer chain with N segments (substrate S) such that site 1 of the enzyme binds selectively segment i of the chain, site 2 binds segment j etc. If the first binding occurs at site 1, the concentration of enzyme-substrate complexes is

$$[ES_1] = [E][S]K_1, \tag{1}$$

where K_1 is the corresponding microscopic binding constant. For a simultaneous binding of segment j to site 2 the chain must form a loop. It is required that the length of the stretched chain between segments i and j is at least equal to the distance h_{12} between sites 1 and 2 at the enzyme:

$$h_{12} \leqslant |j - i| l_0, \tag{2}$$

where l_0 is the segment length. The probability for the formation of such a loop will clearly depend on the flexibility of the polymer chain which may be described by the characteristic ratio $C = \langle r^2 \rangle / N l_0^2$; $\langle r^2 \rangle$ is the mean square end-to-end distance of the unperturbed chain. We may then define an effective bond length $l = C^{1/2} l_0$.

This convention [25, p. 158] has the advantage of letting N be unmodified so that the labelling of polymer segments is not affected. It is equivalent, for our purposes, to another convention where N and l_0 are multiplied by factors C^{-1} and C, respectively, to give the equivalent chain [18,25,26]. Assuming that segment i is fixed at site 1 and that the condition (2) is satisfied, we may then formulate the probability of finding segment j in the neighbourhood of site 2, namely at a distance h_{12} from site 1 and in the solid angle $d\Omega$ pointing towards site 2 [18]:

$$W_j(h_{12}) dh_{12}\, d\Omega = h_{12}^2\, dh_{12}\, d\Omega\, A \frac{\exp(-\alpha_{12}/|j-i|^\gamma)}{|j-i|^{1.5\gamma}}, \tag{3}$$

with $A = (3/2\pi l^2)^{3/2}$ and $\alpha_{12} = 3h_{12}^2/2l^2$.

The exponent γ accounts for excluded-volume effects and is solvent dependent. It will normally lie between 1 and 1.2 [18,25]. In the simplest case, for an ideal solvent, γ is equal to 1. There is another excluded-volume effect due to the presence of the protein. In the limit of very short chains the protein may be considered as a planar absorbing wall, apparently increasing the effective segment length by a factor of 2 [27]. This factor decreases, however, approaching unity as N becomes large, due to the finite dimensions of the protein. For the sake of simplicity we neglect these contributions, especially as detailed knowledge about the shape of the protein would be necessary to include them correctly. This is only a rough approximation for short chains but improves as the chains become longer.

The probability given by eq. (3), divided by the volume element $h_{12}^2 dh_{12} d\Omega$, can be interpreted as the concentration of segment j in the neighbourhood of site 2. Therefore we find for the concentration of complexes with the polymer bound to sites 1 and 2 simultaneously:

$$[ES_{12}] = [ES_1]K_2 c_j = [E][S]K_1 K_2 A \frac{\exp(-\alpha_{12}/|j-i|^\gamma)}{|j-i|^{1.5\gamma}}. \tag{4}$$

Similarly, for simultaneous binding at sites 1, 2 and 3 we obtain

$$[ES_{123}] = [ES_{12}]K_3 c_k = [E][S]K_1 K_2 K_3 \times A^2 \frac{\exp(-\alpha_{12}/|j-i|^\gamma)}{|j-i|^{1.5\gamma}} \frac{\exp(-\alpha_{23}/|k-j|^\gamma)}{|k-j|^{1.5\gamma}}, \tag{5}$$

where $\alpha_{23} = 3h_{23}^2/2l^2$ and h_{23} is the distance between the sites 2 and 3.

In the above treatment we assumed that the selected residues lie on the chain in the order *ijk* or *kji*. If the sequence of selected residues were different, eqs. (3)–(5) could be rewritten with appropriate indexing. It should be noted that once the polymer is bound at three binding sites simultaneously, it is impossible to decide which of the three bindings occurred as

the first, second or third. Clearly it will be necessary to verify for each loop that conditions like eq. (2) are satisfied.

2.2. Nonselective binding

For this case it is assumed that any of the segments in a randomly coiled polymeric substrate may interact with any binding site of the enzyme and that all binding steps occur with identical microscopic binding constants K. It is also assumed that the distance between all binding sites on the enzyme is the same and equal to h. Extensions to more complicated situations should be obvious from the preceding paragraph.

For the binding of one segment to the enzyme we now have:

$$[\mathrm{ES}]^{(1)} = uNK[\mathrm{E}]\,[\mathrm{S}], \tag{6}$$

u being the number of binding sites of the enzyme. If segment i is bound, the concentration of segment j in the neighbourhood of a second enzymatic site is again

$$c_j = A\frac{\exp(-\alpha/|j-i|^{\gamma})}{|j-i|^{1.5\gamma}}, \tag{7}$$

where A and α are defined as in eq. (3) and condition (2) has to be satisfied, with h_{12} replaced by h. The overall concentration of segments in the neighbourhood of the second site on the enzyme will then be

$$c = \sum_{j=i+p_0}^{N} c_j + \sum_{j=1}^{i-p_0} c_j, \tag{8}$$

with p_0 the minimum number of segments which is needed to join two sites by a loop. We take for p_0 the smallest integer that satisfies $p_0 \geqslant h/l_0$.

By averaging over all possible values of i we obtain:

$$\bar{c} = \frac{1}{N}\left(\sum_{i=1}^{N-p_0}\sum_{j=i+p_0}^{N} c_j + \sum_{i=p_0}^{N}\sum_{j=1}^{i-p_0} c_j\right). \tag{9}$$

By inverting the order of summation in the second double sum it may be shown that the two terms in the bracket are equal. This reflects the fact that it is impossible to decide which one of the two bindings occurred first once both sites are occupied.

For the same reason, a statistical factor 1/2 must be included for the substrate as well as for the enzyme. The number of available sites on the enzyme is now $(u-1)$, the corresponding factor $(N-2p_0+1)$ for the substrate is already included in the summation.

The simultaneous binding of two segments of the substrate to the enzyme is thus given by

$$[\mathrm{ES}]^{(2)} = [\mathrm{ES}]^{(1)}K \times \frac{u-1}{2}\frac{A}{N}\sum_{i=1}^{N-p_0}\sum_{j=i+p_0}^{N}\frac{\exp(-\alpha/|j-i|^{\gamma})}{|j-i|^{1.5\gamma}}. \tag{10}$$

Inserting the value of $[\mathrm{ES}]^{(1)}$ and rearranging the summation terms gives

$$[\mathrm{ES}]^{(2)} = \frac{u(u-1)}{2}K^2 \times [\mathrm{E}]\,[\mathrm{S}]\,A\sum_{i=p_0}^{N-1}(N-i)\frac{\exp(-\alpha/i^{\gamma})}{i^{1.5\gamma}} \tag{11}$$

Similarly the binding of three, four, or generally, m segments of a coiled substrate to a single enzyme molecule is calculated to be

$$[\mathrm{ES}]^{(m)} = \binom{u}{m}K^m[\mathrm{E}]\,[\mathrm{S}]A^{m-1}\sum\nolimits_m \tag{12}$$

with

$$\sum\nolimits_m = \sum_{i_1=1}^{N-(m-1)p_0}\sum_{i_2}\cdots \sum_{i_m} g(i_2-i_1)\cdots g(i_m-i_{m-1}), \tag{13}$$

where

$$g(r) = \exp(-\alpha/r^{\gamma})/r^{1.5\gamma}, \tag{14}$$

and the summation over i_j goes from $(i_{j-1}+p_0)$ to $(N-(m-j)p_0)$.

As in eq. (11), the number of summations may be reduced by one. Then:

$$\sum\nolimits_m = \sum_{i_1=(m-1)p_0}^{N-1}(N-i_1) \times \sum_{i_2}\cdots\sum_{i_{m-1}} g(i_2)\cdots g(i_{m-1})g(i_1-i_2-\cdots-i_{m-1}), \tag{15}$$

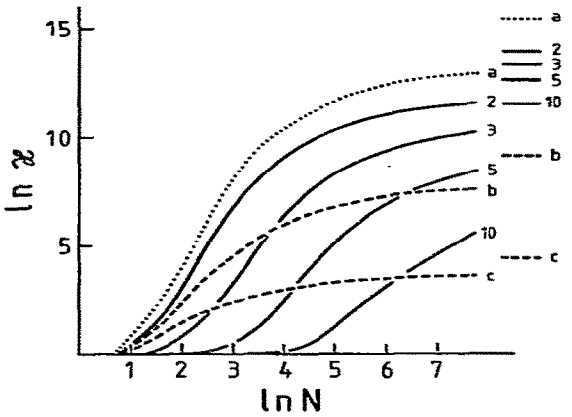

Fig. 1. Double logarithmic plot of κ_1 and κ_2 versus chain length N. The parameters κ_1 and κ_2 are defined by eqs. (19) and (21) respectively: $\kappa_1 - 1$ is the number of complexes with one or more loops and κ_2 is the total number of bound sites both expressed per number of complexes with only one site of the polymer chain attached. Solid lines: κ_1 with $u = 4$ and h/l_0 as indicated. Dotted line (a): κ_2 with $u = 4$ and $h/l_0 = 2$. Dashed lines: κ_1 at fixed $h/l_0 = 2$; (b): $u = 3$; (c): $u = 2$. In all cases $\gamma = 1$ and $l_0 = 3.6$ Å. Asymptotes for $N \to \infty$ are indicated on the right.

the summation over i_j now going from p_0 to $(i_1 - i_2 - \ldots - i_{j-1} - (m-j)p_0)$.

In the case of large N, the calculation of the summations may be simplified by conversion to integrals. Setting $a_0 = p_0/N$ and $x = i/N$, we obtain, for instance, from eq. (11):

$$\sum\nolimits_2 = \sqrt{N} \int_{a_0}^{1} \frac{\exp(-\beta/x^{\gamma})}{x^{1.5\gamma}}(1-x)\mathrm{d}x, \qquad (16)$$

where $\beta = \alpha/N = 3h^2/2CNl_0^2$ depends on h and on the product CN.

In the case of $\gamma = 1$, the integrals may be expanded into error functions, as shown in the Appendix. In this case, the asymptotic limit of Σ_i as N tends towards infinity is

$$\sum\nolimits_m \to N(2/\sqrt{p_0})^{m-1} \qquad (17)$$

Using this formula one may get a rough estimate of the effect of loop formation without going into more complex numerical computations.

2.3. Equilibrium binding and Michaelis constants for the interaction between a polymeric ligand and a multivalent protein

In many cases experimentally, the concentrations of individual complexes cannot be measured but it is possible to determine one of the following quantities:

a) the total concentration of complexes:

$$[ES]^{(1)} + [ES]^{(2)} + \ldots + [ES]^{(u)} = [ES]^{(1)}\kappa_1 = [E_0] - [E], \qquad (18)$$

where $[E_0]$ and $[E]$ are the total and free enzyme concentration, respectively, and κ_1 is defined as

$$\kappa_1 = \sum_i [ES]^{(i)}/[ES]^{(1)}. \qquad (19)$$

b) the total concentration of occupied sites:

$$[ES]^{(1)} + 2[ES]^{(2)} + \ldots + u[ES]^{(u)} = [ES]^{(1)}\kappa_2, \qquad (20)$$

with

$$\kappa_2 = \sum_i i[ES]^{(i)}/[ES]^{(1)}. \qquad (21)$$

For the model described above, both quantities may be determined from the evaluation of eqs. (12)–(15). Because these equations are complex and the contributions of the parameters h and u may not be obvious, κ_1 was evaluated and plotted versus N in a double logarithmic plot (fig. 1) for various values of h and of u. Also shown is a comparison of κ_2 with κ_1.

If the catalytic reaction is preceded by a rapid binding equilibrium of the various $ES^{(i)}$ complexes, a Michaelis constant K_m may be defined. From $\mathrm{d}[P]/\mathrm{d}t = k_{cat}([ES]^{(1)} + 2[ES]^{(2)} + \ldots + u[ES]^{(u)})$ and eq. (18) it follows that

$$\mathrm{d}[P]/\mathrm{d}t = k[E_0][S]/(K_m + [S]) \qquad (22)$$

with

$$K_m = (\kappa_1 uNK)^{-1} \quad \text{and} \quad k = k_{cat}\kappa_2/\kappa_1. \qquad (23,24)$$

K_m and k are the effective constants which may be determined from the slope and the ordinate intercept of a Lineweaver–Burk plot. Contrary to K_m, the turnover number k is only slightly chain length dependent. The extreme case is given for large values of the microscopic binding constant K, where the term

$[ES]^{(u)}$ dominates, where u is the maximum number of binding sites on the enzyme. In this case, the quotient κ_2/κ_1 would tend to u for sufficiently large N. This means that k would only vary by a small number as N goes from 1 to infinity. The large increase of κ_1 with increasing N which gives rise to a strong chain length dependence of K_m will be discussed below for the example of the enzyme prolyl hydroxylase.

3. Chain length dependence of the Michaelis constant of prolyl hydroxylase

Prolyl hydroxylase [13,14] converts approximately 100 proline residues to hydroxyproline in each collagen chain. The enzyme will also react with polypeptides containing prolyl residues in sequences similar to collagen [14]. The smallest peptide which the enzyme recognizes is X-Pro-Gly where X can be any of a variety of amino acids [14]. It is an example of an enzyme which shows an extremely large decrease of its Michaelis constant K_m with the increasing size of the polymeric substrates [14,21–24]. This unusual chain length dependence remains steep even when the number of X-Pro-Gly peptide units exceeds 3 which would correspond to the maximum size of the active site which may be estimated from the geometry of the enzyme [28]. The number of binding sites on the enzyme is not known; however, purified prolyl hydroxylase from chick embryos contains four subunits [29]. Therefore the maximum number of binding sites on the enzyme is probably four.

For prolyl hydroxylase, the Michaelis constants may be treated as equilibrium dissociation constants [24, 30]. The catalytic constant is very small (4 s^{-1}) and essentially independent of the chain length of the substrate [21,24].

The Michaelis constants may now be calculated from eq. (23), where

$$\kappa_1 = 1 + \frac{3}{2N}KA\sum\nolimits_2 + \frac{1}{N}(KA)^2\sum\nolimits_3 + \frac{1}{4N}(KA)^3\sum\nolimits_4 \qquad (25)$$

is given by eqs. (12,19) with $u = 4$.

Whereas the polymer chain consists of N amino acid residues of length $l_0 = 3.6$ Å, each, it is the Pro-Pro-Gly tripeptide units which are recognized and bind to the enzyme. Consequently, summations in eqs. (13) or (15) have to be done in steps of three (i = 1,4,7 ...). For large values of N, and especially when using integrals, it is sufficient to divide by a factor 3 for each summation. Experimental data for the characteristic ratio of poly(Pro-Pro-Gly) do not exist, but C has been found to be approximately 4 for some denatured protein [31]. For the sake of computational simplicity, $C = 3$ was taken as a reasonable assumption. (In this special case another convention can be used for the definition of an equivalent chain, setting $N' = N/3$, $l' = 3l_0$ and $p_0' = p_0/3$. Simple summations can then be used.) It was further assumed that the distance between enzymatic sites is twice the length of one Pro-Pro-Gly unit, or 21.6 Å, and thus $p_0 = 6$. The value of $h = 21.6$ Å may seem small compared to the

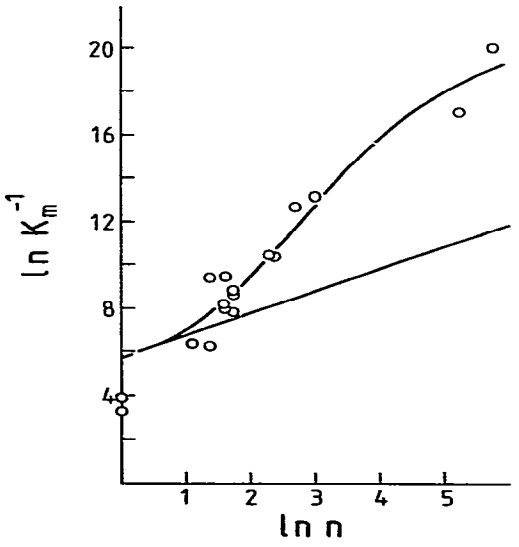

Fig. 2. Double logarithmic plot of K_m^{-1} versus chain length $n = N/3$ (= number of tripeptide units) for polypeptide substrates of prolyl hydroxylase. K_m is the Michaelis constant of the enzyme as measured from the substrate concentration at half maximum velocity. The substrates consist of $(Pro\text{-}Pro\text{-}Gly)_n$ polypeptides where n varies from 1 to 20. Additional substrates include polypeptides containing X-Pro-Gly sequences either in synthetic substrates or in natural substrates such as bradykinin, *Ascaris* cuticle collagen and unhydroxylated embryonic chick collagen [14,24]. The curve through the data points was drawn according to eqs. (12)–(15) with $u = 4$. The straight line of slope 1 would result without loop formation.

diameter of the enzyme which is about 90 Å, but the exact position of the active sites is not known, and a much larger value of h would be incompatible with the strong chain length dependence occurring for very small values of N (fig. 2).

In fig. 2, the experimental data for the binding of prolyl hydroxylase to poly(Pro-Pro-Gly) are given together with a calculated curve using an intrinsic binding constant $K = 85\ M^{-1}$ for tripeptide units. Excluded-volume effects were neglected in the calculation by setting $\gamma = 1$ so that the integrals could be expanded into error functions. The theoretical curve is seen to give a good fit of the data. The two points at the highest N values refer to data obtained on unhydroxylated collagen. They were included in the figure though they are not directly comparable with the poly(Pro-Pro-Gly) results.

4. Discussion

Association equilibria between multisite molecules and independent ligands have been developed to describe interactions between proteins and small molecules [25,32–34]. A treatment, however, has not yet appeared for the situation where the ligands are not independent but occur as segments of a polymer.

In this paper, the concept of ring closure, developed for describing reactions proceeding between segments of a polymer [19,20], was extended to describe the probability of loop formation for a coiled polymer which binds to a multivalent molecule. The sites of the polyvalent molecule were assumed to be separated by finite distances. The mathematical formulation of this mechanism allows the calculation of the equilibria between the various complexes which can be formed by a flexible polymer and a multivalent molecule. The mechanism may be applied to a large number of situations of practical interest. A treatment was first developed for the selective binding of specific segments of a polymer to specific sites of a protein. An example for an application would be an enzyme which may recognize and join two ends of a polymer. Other examples which involve non-selective binding are more common such as binding of poly(Pro-Pro-Gly) to prolyl hydroxylase or the binding of antibodies to a polymeric antigen of repeating determinants [17,35]. A similar mechanism has been invoked by von Hippel et al. to explain the kinetics for the interaction of the lac operon repressor of *E. coli* [36].

The mechanism described above was further developed for the case where all polymer segments and all the binding sites on the protein were respectively equivalent, and that all the interactions had the same intrinsic binding constants. Experimentally observable binding quantities, i.e. the total concentration of complexes and the concentration of occupied sites on the polyvalent protein, were shown to increase abruptly as the chain length N becomes larger than a critical value p_0 and tended to an asymptotic value in the limit of very large N. The parameters used are: the number of binding sites at the protein, u, and their mutual distance, h, the characteristic ratio, C, and segment length, l_0, of the polymer, an excluded-volume exponent γ and the microscopic binding constant K. p_0 is then the smallest multiple of l_0 that equals or exceeds h. The numerical results can be expected to be reasonably accurate for chain lengths that are large compared with the dimensions of the protein. They can be taken as a first approximation for shorter chains where the effects due to the excluded volume of the protein are comparatively important.

A special situation is encountered if the binding of the polymer occurs as a rapid equilibrium step followed by a slow catalytic reaction. Effective Michaelis constants K_m and turnover numbers may then be determined from an observation of the initial rate of product formation. According to our model, a strong chain length dependence of K_m is predicted. On the other hand, the turnover number k_{cat} increases only slightly when N changes from 1 to infinity. These conclusions agree qualitatively with the experimental data in the case of the binding of poly(Pro-Pro-Gly) to the enzyme prolyl hydroxylase [14].

In order to achieve a more quantitative comparison of the model with the experimental data for prolyl hydroxylase, some reasonable ad-hoc assumptions on unknown parameters were necessary. Each enzyme molecule was thought to carry four equivalent catalytic sites at a distance of twice the length of a stretched tripeptide unit. The characteristic ratio C was taken to be 3 for the polymer and γ set equal to unity. With these assumptions, a good fit of the exper-

mental data was obtained, the microscopic binding constant being adjusted to $K = 85\ \mathrm{M}^{-1}$, expressed in reciprocal molarity of substrate (fig. 2). Without loop formation on the enzyme according to our model, the chain-length dependence of K_m^{-1} would have resulted in a straight line of slope 1 as shown in the figure. This is due to the factor N appearing in K_m^{-1} because the Michaelis constants are usually expressed in molarity of polymer. For the same reason, the asymptotic limit of the fitted curve for very large N is also a straight line of slope 1. It was checked that the binding of different substrate molecules to the same enzyme molecule gave negligible contributions for substrate concentrations less than about 3 K_m as used in the experiments.

The statistical treatment given here for the data on prolyl hydroxylase may be less exact for very low values of N for at least three reasons. First, the theoretical curve may only give an upper limit due to the possibility of strong excluded volume effect with the enzyme. Second, an additional chain length dependence due to incomplete filling of a binding site may be applicable (especially for single tripeptides) [16,37]. Third, end effects may also be involved. These complications may help explain why the theoretical curve in fig. 2 does not extrapolate to the value experimentally observed for a single X-Pro-Gly tripeptide.

The mechanism of multiple interaction of a macromolecule with a polymeric ligand offers the first rational explanation of the unusual chain-length dependence of the Michaelis constant of prolyl hydroxylase which is observed in a range in which the chains are already much longer than the possible size of the binding pocket. The formalism presented here is sufficiently general to be applied to many other related processes of biological interest.

Appendix

For large N, summations may be approximated by integrals setting $\mathrm{d}x_i = 1/N$. With the notation of sect. 2, eq. (13) then takes the following form:

$$\frac{1}{N}\sum\nolimits_m = N^{-(m-1)/2} \int_0^{1-(m-1)a_0} \mathrm{d}x_1 \int_{a_2}^{b_2} \mathrm{d}x_2 \cdots \int_{a_m}^{b_m} \mathrm{d}x_m\, g'(x_2 - x_1) \ldots g'(x_m - x_{m-1}), \tag{A1}$$

with

$$a_i = x_{i-1} + a_0, \quad b_i = 1 - (m - i)a_0, \quad g'(y) = \exp(-\beta/y^\gamma)/y^{1.5\gamma}.$$

In the case of $\gamma = 1$, the integrals are of the error function type. After rearranging the order of integration we get the following expressions for Σ_2 to Σ_4. The error function is defined by

$$\phi(z) = \frac{2}{\sqrt{\pi}} \int_0^z \mathrm{e}^{-x^2}\,\mathrm{d}x. \tag{A2}$$

In addition, we define

$$\varphi(z) = \phi(\sqrt{\beta/z}). \tag{A3}$$

Then:

$$\frac{1}{N}\sum\nolimits_2 = (\pi/N\beta)^{1/2}\left\{(1 - a_0)\varphi(a_0) - \int_{a_0}^{1} \mathrm{d}x\,\varphi(x)\right\}, \tag{A4}$$

$$\frac{1}{N}\sum\nolimits_3 = (\pi/N\beta)\left\{(1 - 2a_0)\varphi^2(a_0) - 2\varphi(a_0)\int_{a_0}^{1-a_0} \mathrm{d}x\,\varphi(x) + \int_{a_0}^{1-a_0} \mathrm{d}x\,\varphi(x)\varphi(1 - x)\right\}. \tag{A5}$$

$$\frac{1}{N}\sum\nolimits_4 = (\pi/N\beta)^{3/2}\left\{(1-3a_0)\varphi^3(a_0) - 3\varphi^2(a_0)\int_{a_0}^{1-2a_0} dx\,\varphi(x) + 2\varphi(a_0)\int_{a_0}^{1-2a_0} dx\,\varphi(x)\varphi(1-x-a_0)\right.$$

$$\left. + 2(\beta/\pi)^{1/2}\int_{a_0}^{1-2a_0} dx\,\varphi(x)\int_{(1-x-a_0)^{-1/2}}^{a_0^{-1/2}} dy\,\exp(-\beta y^2)\varphi(1-x-y^{-2})\right\}. \qquad \text{(A6)}$$

Note that for sufficiently large values of $\beta/a_0 = \alpha/p_0$ the factor $\varphi(a_0)$ may be set equal to unity.

By using a "rational" approximation for $\phi(z)$ [38], these integrals are easily evaluated by a numerical procedure, since even in the case of triple binding there remain only single integrals, whereas one double integral has to be calculated in the case of $[ES]^{(4)}$. In the limit of very large N, $\beta = \alpha N^{-1}$ tends to zero and the exponential in the integrals becomes unity. The integrals may then be solved analytically.

For $\gamma = 1$, the leading term of each integral is evaluated to be $2\,(N/p_0)^{1/2}$. Thus

$$\sum\nolimits_m \to N(2/\sqrt{p_0})^{m-1} \qquad \text{(A7)}$$

as N goes to infinity.

References

[1] T. Kornberg and M.L. Gefter, J. Biol. Chem. 247 (1972) 5369.
[2] W.R. McClure and T.M. Jovin, J. Biol. Chem. 250 (1975) 4073.
[3] L.A. Sherman and M.L. Gefter, J. Mol. Biol. 103 (1976) 61.
[4] T.M. Jovin, Ann. Rev. Biochem. 45 (1976) 889.
[5] P. Chambon, in: The enzymes, X, ed. P.D. Boyer (Academic Press, New York, 1974) p. 261.
[6] M.J. Chamberlin, in: The enzymes, X, ed. P.D. Boyer (Academic Press, New York, 1974) p. 333.
[7] N.G. Nossal and M.F. Singer, J. Biol. Chem. 243 (1968) 913.
[8] W. Stalmans and H.G. Hers, in: The enzymes, IX, ed. P.D. Boyer (Academic Press, New York, 1974) p. 167.
[9] S.J. Kerr and E. Borek, in: The enzymes, IX, ed. P.D. Boyer (Academic Press, New York, 1974) p. 167.
[10] R.G. Spiro, M.J. Spiro and A.M. Adamany, Biochem. Soc. Symp. 40 (1974) 37.
[11] J. Molnar, Mol. and Cell. Biochem. 6 (1975) 3.
[12] K.I. Kivirikko and R. Myllylä, Int. Rev. of Connect. Tissue Res., in press.
[13] G.J. Cardinale and S. Udenfriend, Adv. Enzymol. 41 (1974) 245.
[14] D.J. Prockop, R.A. Berg, K.I. Kivirikko and J. Uitto, in: Biochem. of collagen, eds. G.N. Ramachandran and A.H. Reddi (Plenum Publ. Corp., New York, 1976) p. 163.
[15] T.M. Turpeenniemi, U. Puistola, H. Anttinen and K.I. Kivirikko, Biochim. Biophys. Acta 483 (1977) 215.
[16] J. Cisar, E.A. Kabat, M.M. Dorner and J. Liao, J. Exp. Med. 142 (1975) 435.
[17] J.-C. Jaton, H. Huser, W.F. Riesen, J. Schlessinger and D. Givol, J. Immunol. 116 (1976) 1363.
[18] H. Morawetz, Macromolecules in solution (2nd ed., John Wiley and Sons, Inc., New York, 1975).
[19] N. Goodman and H. Morawetz, J. Polymer Sci. 31 (1970) 177.
[20] M. Sisido, Polymer Journal 3 (1972) 84.
[21] R.A. Berg, Y. Kishida, S. Sakakibara and D.J. Prockop, Biochemistry 16 (1977) 1615.
[22] J.O.D. McGee, R.E. Rhoads and S. Udenfriend, Arch. Biochem. Biophys. 144 (1971) 343.
[23] K.I. Kivirikko, Y. Kishida, S. Sakakibara and D.J. Prockop, Biochim. Biophys. Acta 271 (1972) 347.
[24] R.A. Berg and D.J. Prockop, Biochemistry 12 (1973) 3395.
[25] C. Tanford, Physical chemistry of macromolecules (John Wiley and Sons, Inc., New York, 1975).
[26] P.J. Flory, Statistical mechanics of chain molecules (John Wiley and Sons, Inc., New York, 1969).
[27] E.A. Di Marzio, J. Chem. Phys. 42 (1965) 2101.
[28] B.R. Olsen, R.A. Berg, K.I. Kiviriko and D.J. Prockop, Eur. J. Biochem. 35 (1973) 135.
[29] R.A. Berg and D.J. Prockop, J. Biol. Chem. 248 (1973) 1175.
[30] K. Juva and D.J. Prockop, J. Biol. Chem. 244 (1969) 6486.
[31] S. Lapanje and C. Tanford, J. Amer. Chem. Soc. 89 (1967) 5030.

[32] I.M. Klotz, D.W. Darnall, N.R. Langerman, in: The proteins, I, eds. H. Neurath and R.L. Hill (Academic Press, New York, 1975) p. 294.
[33] J.T. Edsall and J. Wyman, Biophysical chemistry (Academic Press, New York, 1958).
[34] D.E. Koshland, Jr., in: The enzymes, I, ed. P.D. Boyer (Academic Press, New York, 1970, 3rd Ed.) p. 342.
[35] H. Maeda, A. Schmidt-Kessen, J. Engel and J.-C. Jaton, Biochemistry 16 (1977) 4086.
[36] P.H. v. Hippel, A. Revzin, C.A. Gross and A.C. Wang, in: Protein–ligand interactions, eds. H. Sund and G. Blauer (de Gruyter, Berlin, 1975) p. 270.
[37] A. Berger and I. Schechter, Phil. Trans. Roy. Soc. London Sec. B 227 (1970) 349.
[38] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions (Dover,Publ., New York, 1965).