The EMBO Journal vol.8 no.3 pp.797 – 804, 1989

SPKK, a new nucleic acid-binding unit of protein found in histone

Masashi Suzuki

Laboratory of Neurochemistry, Division of Molecular Physiology, National Institute for Physiological Sciences, Okazaki 444, Japan

Communicated by A.Klug

A new DNA-binding unit of a protein different from the α-helix, the β-sheet and the Zn-finger is proposed based on the analysis of the structure of the N-terminus of sea urchin spermatogenous histone H1. DNA-binding arms of the sea urchin spermatogenous histones, H1 and H2B, are composed of repeats of Ser-Pro-Lys(Arg)-Lys(Arg) (SPKK) residues. A six-times repeat of SPKK (S6 peptide) was isolated from H1 and the competition of S6 for DNA binding with a DNA-binding dye, Hoechst 33258, was analysed. The S6 peptide is shown to be a competitive inhibitor of Hoechst 33258, and it is concluded that the SPKK repeat binds to DNA in its minor groove with a binding constant, $K_{S6} = 1.67 \times 10^{10}$ M^{-1}. The circular dichroism (CD) spectrum of a synthetic peptide, SPRKSPRK (S2 peptide), is quite different from those of both the α-helix and the β-sheet and resembles that of a random coil. From statistical consideration of protein structures it is proposed that SPKK forms a compact β-turn stabilized by an additional hydrogen bond. Since a repeated chain of such turn of SPKK offers a repeat of amides of Ser residues at a distance similar to that of DNA-binding amides of the drugs, Hoechst 33258 and netropsin, and since the amides of these drugs bind to DNA replacing the spine of hydration in a minor groove, it is proposed that a repeat of SPKK binds to DNA in the minor groove using similar hydrogen bonds.

Key words: Asx-turn/β-turn/DNA binding/histone/sea urchin

Introduction

To date three different DNA-binding units have been proposed. One is the α-helix represented by the helix-turn-helix structure of prokaryotic repressor (Takeda *et al.*, 1983; Anderson *et al.*, 1987), catabolite activator protein (McKay and Steitz, 1981) and eukaryotic homeotic gene products (Laughon and Scott, 1984). Another is the antiparallel β-sheet found in G5BP (Brayer and McPherson, 1983) and DBP II (Tanaka *et al.*, 1984). The other is the Zn-finger originally suggested for transcription factor IIIA (Miller *et al.*, 1985) and found in yeast gene activator proteins (Hartshorne *et al.*, 1986), segmentation gene products (Rosenberg *et al.*, 1986) and steroid receptors (Green *et al.*, 1986; Green and Chambon, 1987; Godowski *et al.*, 1987; Sap *et al.*, 1986).

A quite new mode of DNA binding is expected in sea urchin spermatogenous histones, H1 and H2B. Primary structures of sea urchin spermatogenous histones, H1 and H2B, were determined by von Holt *et al.* (1984), who pointed out that there was a repeating unit in the N-terminal regions composed of a pentapeptide starting with Pro. Poccia (1987) suggested that it was very likely that these N-terminal regions were made up of β-turns. But the unit of β-turn was not uniquely specified and even Lys-Lys-Ser-Pro (KKSP) was nominated as one of the candidates. The latter, however, cannot form the β-turn, because Pro, in the $i + 3$-th position, lacks NH as a potential hydrogen bonding donor. I have proposed that it is Ser-Pro-Lys-Lys (SPKK) which is the unique repeating unit in N-termini of sea urchin spermatogenous histones, H1 and H2B. This can be seen by comparing both regions carefully (Suzuki, 1988a). If both N-termini are aligned from this viewpoint, they are essentially the same. von Holt *et al.* (1984) stated that the C-terminus of H1 was composed of three repeats of 11 residues and four repeats of more incomplete nature, but I found six or seven simple repeats, including the KKSPKKA sequence, there (Suzuki, 1988b). I concluded the most important part in this repeat is not KKS(T)P but S(T)PKK by comparing primary structures of H1s of other species (Suzuki, 1988b).

The structure of H2B is composed of two domains, the N-terminal elongated DNA-binding arm and the remaining hydrophobic core which interacts with other histones to form the core of the nucleosome (Zalenskaya *et al.*, 1981). The structure of H1 is composed of three domains called nose, head and tail (Hartman *et al.*, 1977). The nose and tail form two elongated DNA-binding arms, while the head is a hydrophobic core. The N-termini of both histones are composed of five to six repeats of SPKK residues (Suzuki, 1988a), while the C-terminus of sea urchin spermatogenous H1 is composed of six or seven repeats, including SPKK (Suzuki, 1988b). Therefore, sea urchin spermatogenous H2B has a six-times repeat of SPKK in its N-terminus, while H1 has six in the N-terminus and five in the C-terminus.

How might such a SPKK repeat bind to DNA? An ordered zig-zag string of water molecules down the minor groove of the DNA, stabilizing the B-form, was found and called 'the spine of hydration' by Kopka *et al.* (1983). This is composed of two layers of water molecules. The first layer connects hydrogen-bonding acceptors of adjacent base pairs in the minor groove of B-DNA, and the second layer connects adjacent water molecules of the first layer. A mode of DNA binding of drugs related to this spine was found and it is quite different from the modes of DNA binding of proteins proposed up till now (Kopka *et al.*, 1987). Two DNA-binding drugs, netropsin and Hoechst 33258, have been co-crystallized with DNA and analysed. In both cases drugs were located in the minor groove of DNA, replacing water molecules of the first layer by amides and the second layer by a covalently connected backbone chain. An amide binds to every two hydrogen acceptors of adjacent bases of

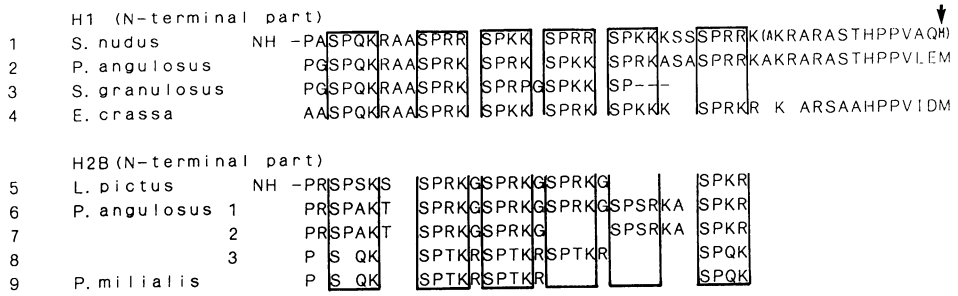

Fig. 1. Primary structure of the isolated N-terminus of H1, S6 peptide, compared with N-termini of spermatogenous H1s and H2Bs of other species of sea urchin. The primary structure of the N-terminus of spermatogenous histone H1 of the sea urchin, *S.nudus* (**1**), was determined by a protein sequencer and that of the isolated peptide was verified by an amino acid composition analysis. Sequences in parentheses were determined from the amino acid composition and comparison with other sequences. An arrow shows the cleaved methionine. N-termini of sea urchin spermatogenous H1 (**2**–**4**) and H2B (**5**–**9**) are essentially composed of six repeats of Ser-Pro-Lys-Lys (SPKK) as boxed. These sequences were derived from the papers of von Holt *et al.* (1984), Lai and Childs (1986) and Busslinger and Barberis (1985). Arg residues are not distinguished from Lys residues and a little deviation from the standard is permitted.

different chains and forms a three-centred hydrogen bonding, so that the distance of neighbouring amides of the drugs corresponds to that of adjacent base pairs.

Since the binding of Hoechst 33258 to DNA can be measured by fluorescence, and since the mode of DNA binding of Hoechst 33258 and netropsin has been determined at atomic resolution, it was decided to compare the DNA binding of these drugs and that of the SPKK repeat, to try to reveal the mode of DNA binding of SPKK.

Results and discussion

The primary structure of the isolated peptide

N-termini of H1 were isolated by CNBr cleavage following chromatography using a CM-Toyopearl 650M column (Figure 2a). The primary structure of the peptide, S6 peptide, isolated from spermatogenous H1 of the sea urchin, *Strongylocentrotus nudus*, is shown in Figure 1. This is composed of a six-times repeat of SPKK followed by additional 16 residues, including four basic residues, and thus resembles the corresponding parts of H1s of other species of sea urchin already reported. Binding of S6 to DNA was examined by DNA affinity chromatography (Figure 2b). S6 peptide was eluted at 0.34 M NaCl and thus shown to bind strongly to DNA.

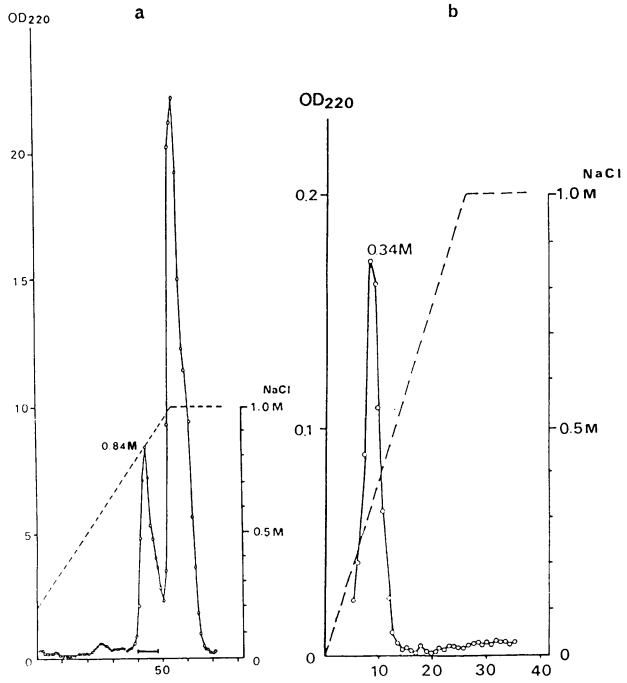

Fig. 2. Elution profiles of the S6 peptide. (**a**) Elution profile of the CM-Toyopearl 650 M column. H1 was cleaved by CNBr and loaded on to a CM-Toyopearl 650 M column pre-equilibrated with 20 mM Na-phosphate buffer, pH 6.2. The column was eluted with a linear gradient of NaCl from 0 to 1 M in 20 mM Na-phosphate buffer, pH 6.2. Two peaks were observed, the lower and higher peaks corresponding to the N-terminus, S6 peptide and C-terminus peptide respectively. Fractions shown by a thick line were combined and used for experiments. (**b**) Elution profile of the DNA cellulose column. DNA binding of the S6 peptide was examined by DNA affinity chromatogrpahy. The DNA cellulose column was pre-equilibrated with 10 mM Tris–HCl buffer, pH 8.0, and it was eluted with a linear gradient of NaCl from 0 to 1 M in 10 mM Tris–HCl buffer, pH 8.0. S6 peptide was eluted at 0.34 M and thus proved to bind strongly to DNA.

Competition of SPKK with Hoechst 33258 for DNA binding

The dependence of the fluorescence intensity on the concentration of dye in the absence of S6 peptide was measured as shown in Figure 3a. Inhibition of DNA binding of Hoechst 33258 by S6 peptide was measured at different concentrations of S6 peptide, in the range 2.07×10^{-9} to 3.30×10^{-8} M, and Hoechst 33258, in the range 5.34×10^{-9} to 4.27×10^{-8} M (Figure 3b). As the concentration of S6 peptide was increased, the intensity of the fluorescence decreased remarkably.

The mode of inhibition of binding of the dye to DNA by the peptide was analysed using the same method as for determining the mode of inhibition of binding of a substrate to an enzyme by an inhibitor. The DNA binding constant of the dye then corresponds to the Michaelis constant. Since the intensity of fluorescence, F, corresponds to the amount of dye bound to DNA, [DNA–dye], the mode of inhibition by S6 peptide to DNA binding of the dye was analysed by Lineweaver–Burk plot (Figure 4a), a plot of $1/F$ versus 1/[dye], where [dye] is the total concentration of added dye. The Lineweaver–Burk plot shows a plot of typical competitive inhibition at concentrations of dye $< 2.14 \times 10^{-8}$ M, because all of the three lines show the same $1/F$ intercepts.

However, at a dye concentration of 4.27×10^{-8} M,

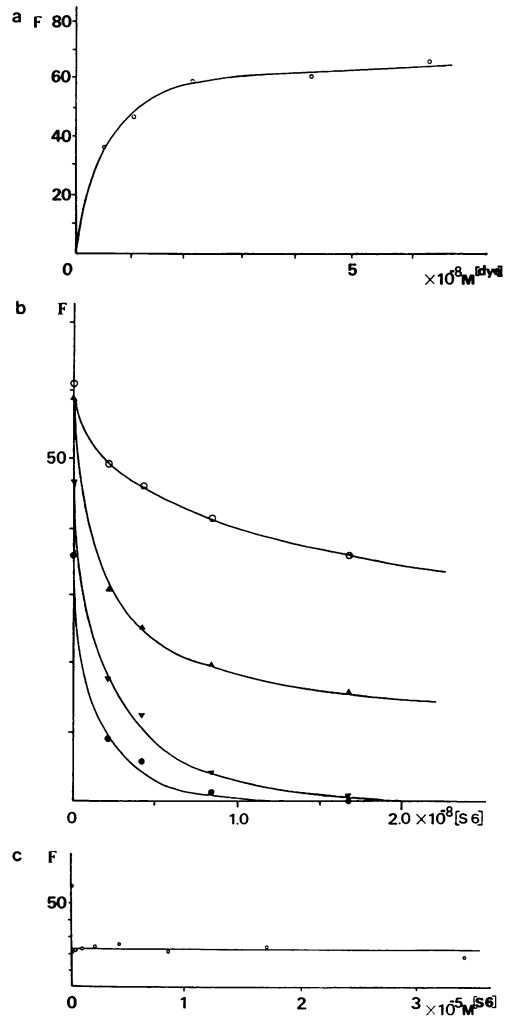

Fig. 3. Inhibition of DNA binding of Hoechst 33258 by S6 peptide. (**a**) Dependence of fluorescence intensity on the concentration of Hoechst 33258. Fluorescence from the DNA–Hoechst complex in the absence of S6 peptide was measured at 450 nm, excited at 354 nm. The concentration of DNA was fixed to 8.08×10^{-8} M and the concentration of Hoechst 33258 was changed from 5.34×10^{-9} to 6.4×10^{-8} M. The difference in intensities of the solution and the dye itself at the same concentration was calculated and shown as real intensity of fluorescence, F, in the vertical axis. Thus the value of F corresponds to the concentration of the DNA–dye complex. (**b**) Dependence of fluorescence intensity on the concentration of S6 peptide. The concentration of DNA was fixed at 8.08×10^{-8} M. At the concentration of Hoechst 33258, 5.34×10^{-9} M (●), 1.07×10^{-8} M (▼), 2.14×10^{-8} (▲) and 4.27×10^{-8} M (○), the concentration of S6 peptide was changed from 2.07×10^{-9} M to 3.30×10^{-8} M. The remarkable decrease of the fluorescence intensity was observed as the increase of the concentration of S6 peptide. (**c**) Dependence of fluorescence on the concentration of the S6 peptide at high concentration of the dye. The concentration of DNA and Hoechst 33258 were fixed at 8.08×10^{-8} and 4.27×10^{-8} M respectively. The concentration of S6 peptide was increased up to 3.4×10^{-5} M, but the intensity of fluorescence did not decrease from the initial level. In this communication all of the concentrations of DNA are expressed as those of phosphates.

deviations from the linearity can be seen especially at higher concentrations of S6 peptide. This can be explained by the fact that the Hoechst dye shows a secondary mode of DNA binding, in which the drug molecule lies outside the grooves on the surface of the helix (Bontemps *et al.*, 1975). The binding constant of the secondary mode is smaller than the first mode by one or two orders of magnitude and is highly cooperative (Bontemps *et al.*, 1975). The deviations at higher concentrations of both the dye and the peptide can be accounted for by this weaker second DNA-binding mode: the peptide inhibits only the DNA binding by the dye in the first mode. At high concentrations of the peptide, binding of the dye in the second mode begins to occur, because the dye cannot bind to the sites of the first mode of binding already occupied by peptides. Thus when the concentration of Hoechst 33258 was low, 5.34×10^{-9} M and 1.07×10^{-8} M, fluorescence was completely extinguished by the addition of S6 (Figure 3b). However, when the concentration of Hoechst 33258 was high, 7.16×10^{-8} M, fluorescence could not be decreased below a certain intensity even in the presence of S6 peptide as high as 3.4×10^{-5} M (Figure 3c). This intensity comes from the dye binding to DNA in the second mode.

From the intercept on 1/[dye] axis of the line in the Lineweaver–Burk plot in the absence of the peptide, the binding constant of the dye, K_d, was found to be 2.34×10^8 M^{-1}. The same intercept of the line in the presence of the peptide gives the value of $K_d/(1 + K_{app}[S6])$, where K_{app} is the apparent binding constant of the peptide in the presence of the dye and [S6] is the concentration of the peptide. $1/K_{app}$ was calculated at the four different concentrations of the peptide and plotted versus 1/[S6] (Figure 4b). From this plot the real binding constant of the peptide, K_{S6}, was deduced as 1.67×10^{10} M^{-1}. Since the free energy, ΔG, is calculated as RT In K, ΔG of the Hoechst 33258 and the peptide at 25°C are 11.4 Kcal/mol and 13.7 Kcal/mol respectively. The value for Hoechst 33258, 11.4 Kcal/mol, seems quite reasonable, compared with that of netropsin, which has five amides binding to DNA, 12.2–12.7 Kcal/mol (Kopka *et al.*, 1987), because Hoechst 33258 has two such amides in addition to one pieperazine ring and a hydroxyl group which also may contribute to DNA binding. Bontemps *et al.* (1975) reported the binding constant of Hoechst as 10^6 in 0.5 mM phosphate buffer, pH 7.0, and this extremely low ionic strength may be one reason why the value was so small compared with the value I obtained at an ionic strength close to physiological. I did not repeat the experiment with 0.5 mM phosphate buffer, because at such low ionic strength the electric potential of the dialysis membrane used by these authors cannot be ignored for the calculation of the binding constant and, furthermore, the double helix of DNA becomes very unstable because of the repulsion of two phosphate helices. The value for the S6 peptide, 13.7 Kcal/mol, is rather high and the equivalent of more than five hydrogen bonds is expected to be formed between S6 and DNA.

Mode of binding of SPKK to DNA

From these results, it is concluded that the S6 peptide is a competitive inhibitor of Hoechst 33258 for DNA binding in the first mode, which is found in the co-crystal of Hoechst 33258 and CGCGAATTCGCG, but not that of the second mode. This means the mode of binding of the S6 peptide may well resemble minor groove binding.

One formal possibility for binding would be intercalation, but this seems highly improbable, because in SPKK there are no residues like Tyr which have side chains capable of intercalating between base pairs. Unwinding of DNA would also inhibit the binding of Hoechst but this seems highly unlikely because both CD and NMR spectra of the complex

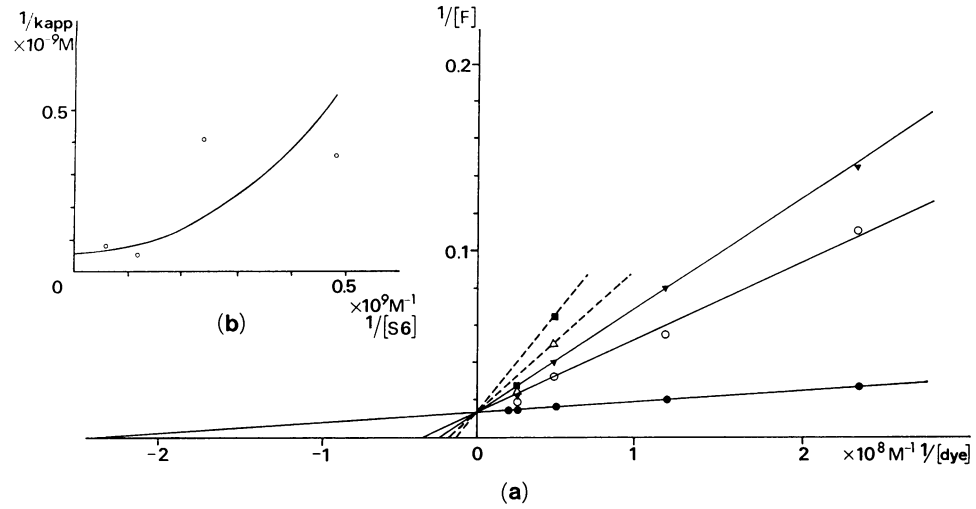

Fig. 4. Lineweaver–Burk plot of DNA binding of Hoechst 33258 in the presence of the S6 peptide. (**a**) The Lineweaver–Burk plot, a plot of $1/F$ versus 1/[dye] where F and [dye] are the fluorescence intensity and the concentration of whole dye respectively, shows the typical characteristics of competitive inhibition. The concentrations of the Hoechst dye used are 5.3×10^{-9} M (●), 1.07×10^{-8} (▼), 2.14×10^{-8} M (▲) and 4.27×10^{-8} M (○) respectively. As shown by broken lines at higher concentrations of both the dye and the S6 peptide the intensities of fluroescences were higher than values expected from values at lower concentrations shown by broken lines. This is explained by the second DNA-binding mode of the dye which is not inhibited by the peptide (see text). Therefore, it is concluded that the S6 peptide binds to the minor groove like Hoechst 33258 of the first DNA-binding mode. From the 1/[dye] intercept binding constant of the dye, K_d, was calculated as 2.34×10^8 M^{-1}. (**b**) Binding constant of the S6 peptide. Binding constants of the appearance of S6 peptide in the presence of the Hoechst dye, K_{app}, was obtained from 1/[dye]intercepts, $1/K_{app}$ was plotted versus 1/[S6], and from $1/K_{app}$ intercept, where [S6] is infinite and therefore the presence of the dye is ignored, the binding constant of S6 peptide is determined as 1.67×10^{10} M^{-1}.

of the synthetic double repeat of SPKK, which I call S2, and the oligonucleotide, CGCATGCG, still showed the characteristics of B-DNA (unpublished data).

It is of course possible that the peptide binds to the major groove and that changes in the structure of the double helix then weakens the binding of Hoechst on the opposite side in the minor groove. However, various experiments with a dimer of the repeating unit, S2, support the hypothesis of binding in the minor groove.

Firstly, the chemical shift of the 1'H proton of a deoxyribose of the above oligonucleotide, which is directed to the minor groove, was changed by the binding of the peptide (unpublished data). Secondly, chemical footprinting experiments of a complex of S2 with a long DNA fragment of 169 bp show protections of adenine bases against methylations (M.Suzuki, unpublished data) and also protections of AT-rich sequences against hydroxyl radical footprinting (M.Churchill and M.Suzuki, in preparation). Since the N3 of adenine, which is the site of methylation, lies in the minor groove and since the hydroxyl radical attacks from the minor groove direction (Tullius *et al.*, 1987), these experiments demonstrate that SPKK-containing peptides do indeed bind in the minor groove. Moreover, the fact that the AT-rich sequences protected from hydroxyl radical cleavage coincide with some of those protected in similar experiments using a complex of the Hoechst dye with the same DNA fragment speaks very strongly for an actual mode of binding which is very similar to that of Hoechst 33258 (M.Churchill and M.Suzuki, in preparation).

Structure of SPKK

The CD spectrum of Ser-Pro-Arg-Lys-Ser-Pro-Arg-Lys, S2 peptide, showed negative ellipticity at 198 nm and no positive ellipticity (Figure 5). This feature is not found at all in the case of both the α-helix and the β-sheet, the spectra of both of which show negative peaks of ellipticity at wave

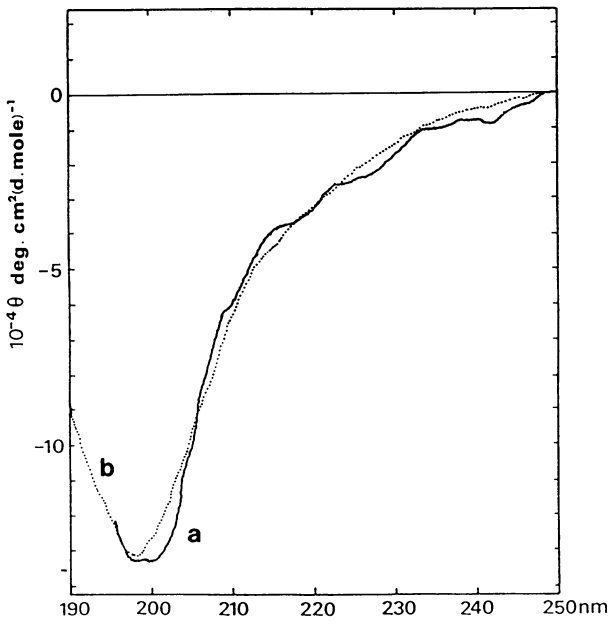

Fig. 5. CD spectrum of SPRKSPRK, the S2 peptide. CD of the octapeptide was measured at the concentration of 2×10^{-2} mg/ml in 1 mM Na-phosphate buffer containing 0.15 M NaCl (a) and H_2O (b). The observed CD spectra show characteristics of random coil but neither that of α-helix nor β-sheet.

lengths from 200 to 230 nm and positive ellipticity at wave lengths from 190 to 200 nm (Cheng *et al.*, 1972). The CD spectrum of the so-called 'random coil' shows the strong negative ellipticity at near 200 nm (Cheng *et al.*, 1972) and resembles the observed spectrum. However, the term 'random coil' here means only that the structure is different from the α-helix and the β-sheet.

The number of the amino acid residues of the unit, four, is that of the β-turn, and the result of secondary structure

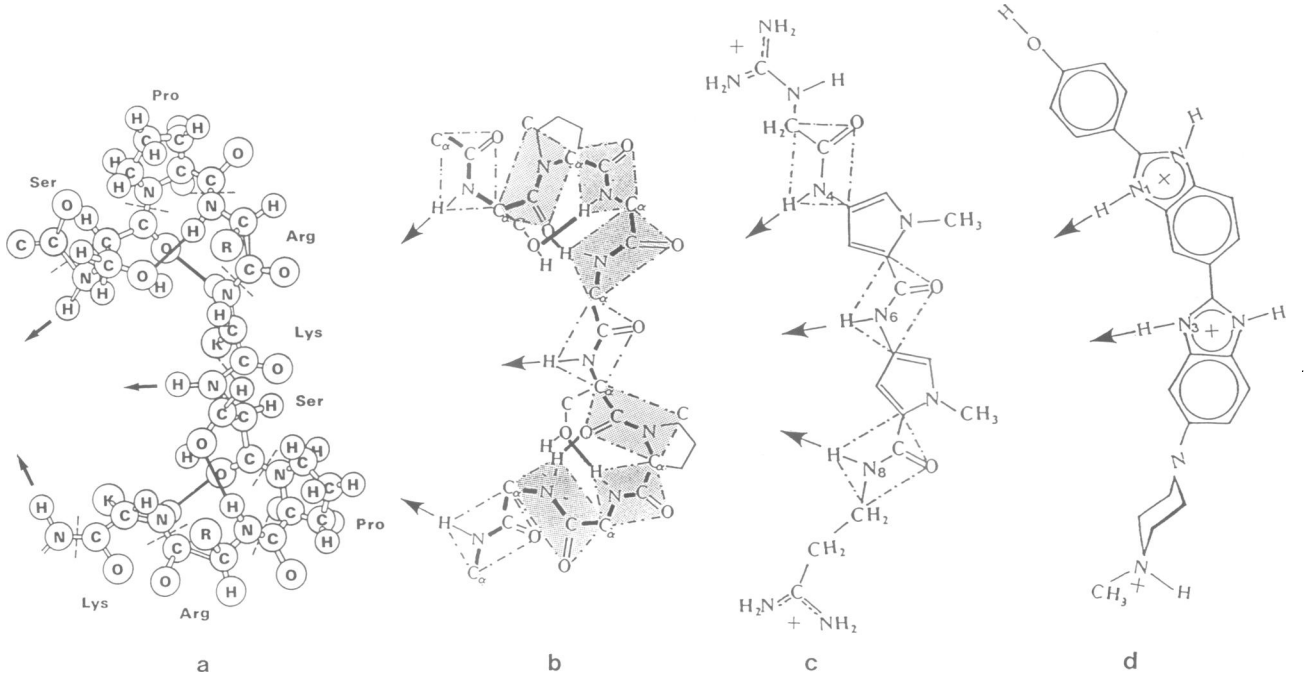

Fig. 6. The structure of SPRK compared with those of DNA-binding drugs. (**a**) Turn structure of SPRK. The proposed structure for SPRK is made by a combination of a β-turn and an Asx turn. Thick solid lines show hydrogen bonds. Two hydrogen bonds stabilize the turn structure. One is formed between the i-th side chain and the $i+2$-th amide. The other is formed between the i-th carboxyl and the $i+3$-th amide. (**b**)–(**d**) Structure of a repeat of SPRK (**b**) compared with those of the DNA-binding drugs, netropsin (**c**) and Hoechst 33258 (**d**). DNA-binding drugs have amides which bind to DNA by replacing the spine of hydration and form three centred hydrogen bond with adjacent base pairs. The SPRK repeats have amides of Ser residues not involved in hydrogen bonding.

prediction, using the method of Chou and Fasman (1974), suggests that SPKK forms a β-turn using Ser at the first position (Suzuki, 1988a). This may be due to the fact that among the 20 amino acid residues, Pro has the highest tendency to occupy the second position in the turn, and that Ser is one of the amino acid residues that is found most frequently at the first position of the turn. Lewis *et al.* (1973) have pointed out that slightly hydrophilic residues such as Ser facilitate turn formation.

These statistical suggestions are explained by Baker and Hubbard's (1984) collection of hydrogen bonds found in proteins. This shows that there are additional hydrogen bonds which stabilize the β-turn in proteins. One of them is a bond between the OH of the i-th residue, which might be Ser, Thr or Asx, and the NH of the $i+2$-th residue (Figure 6a and b). The side chain–main chain interaction of this type can produce a turn by itself, even without the usual i-th CO–$i+3$-th NH bond of the β-turn, and called an 'Asx turn' by Rees *et al.* (1983). Alber *et al.* (1987) made a series of changes in lysozyme in which the amino acid originally at the i-th position of an Asx turn was replaced by other amino acids, sequentially. They analysed the structural changes at atomic resolution and compared the stabilities of the changed proteins. Asp, Asn, Ser and Thr form a similar Asx turn and the stabilities of artificial proteins having these residues were similarly high.

Therefore the unit composed of SPKK can make a turn stabilized by two hydrogen bonds, between the i-th OH of Ser and the $i+2$-th NH, and between the i-th CO and the $i+3$-th NH. This structure is a combination of an Asx turn and a β-turn using the same Ser residue at the i-th position (Figure 6a and b). Such turns with two hydrogen bonds are quite often found in protein structures and examples are listed by Baker and Hubbard (1984) and also described in the recent work of Wilmot and Thornton (1989).

Variations in the repeating unit can be seen in the first unit of the N-termini of H1 and more often in that of H2B (Figure 1). These units may form turns in the sequence of SPXK with a hydrophilic amino acid residue or Ala as X. It should be noted that almost all the $i+3$-th positions are occupied by Lys residues while bulky residues such as Arg, Gln and Thr are found at the $i+2$-th positions.

Comparison of structures of SPKK and netropsin

Since a repeat of SPKK is a competitive inhibitor to Hoechst 33258 for DNA binding and binds to DNA in a similar manner, and moreover since SPKK may form a turn, it is expected that the two molecules, Hoechst 33258 and a turn composed of SPKK, should have some resemblance in their structures, especially in DNA-binding sites. As mentioned in the Introduction, Hoechst 33258 and netropsin bind to DNA by using their amides to form three-centred hydrogen bonds (Figure 6c and d). In netropsin three of these amides are in the planes of peptide bonds. There is only one amide free for hydrogen bonding in one SPKK unit, an amide of Ser, because the other two amides of basic residues form intra-unit hydrogen bonds and Pro lacks an amide (Figure 6a and b). Therefore I have devised a model in which the free amides in a repeating structure of SPKK units were overlapped on the amides of netropsin. The structure proposed above for SPKK, namely a combination of Asx turn and β-turn, was built using coordinates of β-turn I with a fixed ψ angle of the i-th Ser, equal to that found for the turn made by the sequence Ser-Glu-Thr-Tyr in carboxypeptidase (5CPA; Rees *et al.*, the Brookhaven Protein Data Bank), one example of such turns listed by Baker and

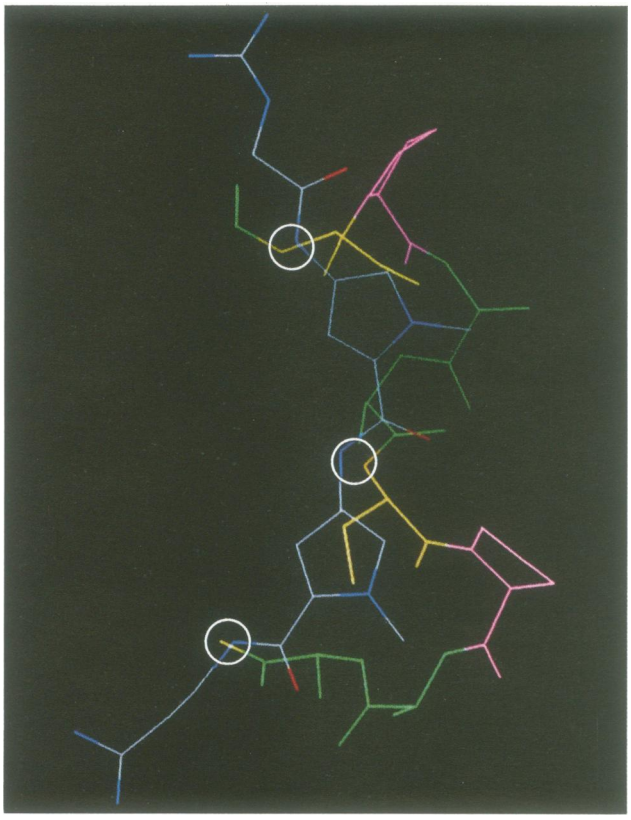

Fig. 7. Comparison of SPRK and netropsin. The model shown is one of a number which were built to study the structural relationship between SPRK and netropsin. The double repeat of SPRK is drawn with different colours for the amino acid residues: yellow for Ser, pink for Pro and green for basic residues. A netropsin molecule is drawn with different colours for the atoms: blue for N, white for C and red for O. A SPRK structure can be built with its amides exactly overlapped onto the nitrogen atoms of netropsin, but the model shown is one in which this condition is relaxed somewhat, because it potentially gives a better fit to a DNA double helix. Side chains of basic residues are omitted except Cβ atoms. Three sets of closely equivalent NHs are indicated by circles.

Hubbard (1984). Successive turns were combined as SPRK−SPRK, with the C-terminus in the form of CONH. Since the essence of the structure lies in the values of the ψ angle of the i-th residue and (ϕ,ψ) of the $i+1$-th and $i+2$-th residues, these angles were fixed and other angles were rotated to bring the three nitrogen atoms of the amides of Ser-1, -5 and the amide of the C-terminal end, which I will call that of Ser-9, into overlap with the three nitrogen atoms of N4, 6 and 8 of netropsin. Since there are three angles involved in making the correspondence, the structure can be formed by variations over a small range of combinations.

With dihedral angles for Ser, Pro, Arg and Lys equal to (−63°,141°), (−60°,−30°), (−90°,0°) and (60°,160°) respectively, a good correspondence between the positions of the nitrogens of the two molecules was observed (Figure 7). The distances between N4, 6, 8 of netropsin and nitrogen atoms of Ser-1, -5, -9 were 0.36, 0.65, 0.35 Å respectively. The distance between two successive amides of Ser residues were respectively 5.6 Å (Ser-1−Ser-5), 5.6 Å (Ser-5−Ser-9), and 9.7 Å (Ser-1−Ser-9), values which seem quite reasonable, since those of netropsin are 5.3 Å (N4−N6), 5.1 Å (N6−N8) and 10.0 Å (N4−N8). The correspondence of the position of amides in both molecules is rather good and the side faces of the model for the repeating SPKK structure, which bears these amides, are very flat, actually even more flat than those of netropsin (Figure 7). Therefore the proposed structure has the simple side faces for putative DNA binding which would not produce any steric hindrance to the side floors of the bases. These characteristics of the SPKK repeat in common with DNA-binding drugs suggest that the details of DNA binding should be similar for the two molecules. Therefore it is very probable that the repeat of SPKK binds to DNA using amides in the same way as amides of Hoechst 33258 and netropsin.

This model can explain why S6 peptide is a competitive inhibitor of the dye. Since both molecules bind to the same structural feature of DNA using the same manner of hydrogen bonding, competitive inhibition should be expected. The binding constant of the S6 peptide, 13.7 Kcal/mol, seems quite reasonable, because S6 should have six NH groups binding to DNA by hydrogen bonds and at least 12, maybe 16, basic residues forming salt bridges with phosphates. The binding constant is expected to be bigger than that of netropsin ~12.4 Kcal/mol, which has only five NH groups binding to DNA. Even if a single SPKK unit forms a turn by itself, it is very probable that this particular arrangement of turns is only stable in the complex with DNA and random coil-like in the absence of DNA, as suggested by the CD spectrum.

According to the proposed model there is one SPKK unit for each base pair and each unit has two basic residues electrically equal to the two phosphates of a base pair. It might be noted that in the case of protamine, Feughelman *et al.* (1955) pointed out that the repeat distance along a fully extended polypeptide chain is equal to the separation of the bases of the nucleotide chain, and the distance between the ends of the successive base side chains pointing in opposite directions is equal to the separation, in the helix axis direction of phosphates on the two helical strands of the nucleic acid molecule.

If we take coordinates of DNA from the complete structure of the netropsin−DNA dodecamer complex (Kopka *et al.*, 1983) and substitute it by the model of SPKK, which simulates the amide of netropsin exactly (cf. Figure 7), there is a general fit but a clash develops between the β-carbon of Lys in the $i+3$-th part and the DNA backbone. It is clear that the detailed structure of the DNA must be changed somewhat and the exact fit to netropsin be relaxed. Model building is in progress.

General discussion

The structure of a nucleosome core particle and a histone octamer was determined by Klug *et al.* by using the methods of diffraction physics (Klug *et al.*, 1980; Finch *et al.*, 1980 Richmond *et al.*, 1984). Another group (Burlingame *et al.*, 1985) also reported the structure of a histone octamer. The structure and location of every histone in the nucleosome can be described using these physical and other biochemical data (Bradbury *et al.*, 1981). Since two molecules of H2B are at the positions of both atria in one heart-shaped nucleosome particle and one H1 molecule is thought to attach to the core particle under the position of the ventricle, one sea urchin spermatogenous nucleosome particle has four arms of SPKK repeats at all four corners near the linker DNA, which is connecting neighbouring nucleosome core particles (Suzuki, 1988b). The sea urchin spermatogenous

nucleosome is different from a typical one in its extraordinarily long (Bradbury *et al.*, 1981) and tightly protected (McGhee *et al.*, 1983) linker DNA besides the bigger H2B and H1 (von Holt *et al.*, 1984). It seems quite reasonable to conclude that the SPKK repeats bind to linker DNA (Suzuki, 1988b). Since in the case of the 5S RNA gene the linker DNA is the starting point for a transcription factor (Rhodes, 1985) and since the linker DNA is also most readily attacked by DNases (McGhee *et al.*, 1983), the repeats may inhibit not only hydrolysis of DNA by DNases but also the transcription of the gene.

In transcription factors S(T)PXX sequences are found on both sides of specific DNA-recognition regions such as Zn-fingers and helix-turn-helices with a frequency higher than that in non-DNA-binding proteins (Suzuki, 1989). This suggests that S(T)PXX sequences contribute to DNA binding in a manner important for gene regulation. Since, as mentioned above, S(T)PXX is enough for forming a β-turn, S(T)PXX is expected to bind to DNA in the same manner as that of SPKK proposed here.

The advantage of a SPKK motif for DNA binding is that the strength of binding to DNA can be modified by the phosphorylation of Ser residues. Actually we have succeeded in isolating the cAMP-dependent kinase which acts on SPKK from the sea urchin embryo and find that the phosphorylation weakens the DNA binding by H1 (M.Suzuki *et al.*, in preparation).

It should be noted that H1s of other species have S(T)PKK sequences in their C-termini (Poccia, 1987; Suzuki, 1988b), while S(T)PXX sequences are frequently found in transcription factors (Suzuki, 1989). Since hydrophilic residues but not basic residues are used at position of X in transcription factors (Suzuki, 1989), the DNA-binding constants of these S(T)PXX arms are smaller than that of S(T)PKK repeats of H1s and thus H1 can inhibit the binding of transcription factors. The mechanisms for removing SPKK arms from DNA for gene activation may be expected to be similar to that found in the fertilization of sea urchins (M.Suzuki *et al.*, in preparation) and there may be a kind of switch for gene regulation using SPKK of H1 and SPXX of transcription factors.

Materials and methods

Sperm of the sea urchin, *S.nudus*, was obtained by injecting solution containing 0.6 M KCl into mature bodies. Histone H1 was extracted from sperm by the acid extraction method with 20% perchloric acid. CNBr cleavage was carried out in 70% acetic acid for 48 h at room temperature. After freeze drying, 100 mg of cleaved H1 was dissolved in 2.5 ml of 20 mM Na-phosphate buffer, pH 6.2, and the solvent was completely exchanged with 20 mM Na-phosphate buffer, pH 6.2, by gel filtration using a PD-10 column (Pharmacia). The fragments of H1 were loaded on to a 100 ml column of CM-Toyopearl 650 M pre-equilibrated with 20 mM Na-phosphate buffer, pH 6.2, and eluted with a linear gradient of NaCl from 0 to 1 M in 400 ml of 20 mM Na-phosphate buffer, pH 6.2. The amino acid composition and primary structure of the isolated fragment was determined by the use of an amino acid composition analyser (Hitachi 835) and a protein sequencer (Applied Biosystem 4704). The concentration of the peptide was determined by measuring the concentration of each amino acid in a certain volume of solution after hydrolysis. DNA affinity chromatography was carried out using 5 ml of a DNA cellulose column (Pharmacia) pre-equilibrated with 10 mM Tris–HCl, pH 8.0. After loading the peptide, the column was eluted with a linear gradient, 0–1 M, of NaCl in 30 ml of 10 mM Tris–HCl, pH 8.0.

Fluorescence of the solutions containing 10 mM Na-phosphate buffer pH 7.0, 150 mM KCl, a fixed concentration of DNA, 8.08×10^{-8} mM in phosphate concentration, and various concentrations of the S6 peptide and a dye, Hoechst 33258, were measured at 450 nm exited at 354 nm using a fluorescence spectrophotometer (Hitachi MPF-4). The difference of intensity of fluorescence of the solution from that containing only the same amount of dye was calculated as true fluorescence, *F*. Hoechst 33258 and DNA of salmon sperm were purchased from Wako Chemical and Sigma Co. respectively. In this communication all of the concentrations of DNA were expressed as those of phosphates.

The octapeptide, SPRKSPRK (S2), was a gift from Applied Biosystem, Japan. The S2 peptide was dissolved in a 1 mM Na-phosphate buffer (pH 7.0) containing 0.15 M NaCl or distilled water to the concentration of 2×10^{-2} mg/ml. Circular dichroism of the S2 peptide was measured using a dichrograph (Jasco A-3020).

A computer graphics study was carried out using a graphic display (COMTEC DS 371) controlled by a computer, A-70 (PFU), using a software, Mol Graph (version 1,0), developed by Daikin Industrial Ltd (Osaka). Coordinates of netropsin with DNA, CGCGAATTCGCG, was according to 6BNA (H.Drew and R.Dickerson) in PDB (Brookhaven).

Acknowledgements

I thank Professor Ebashi for giving me a chance to start this work in his laboratory, Miss H.Kajiura, Mr H.Hattori and Mrs M.Sugiura for technical assistance, Drs M.Fukushima, N.Yagi, K.Mogami, I.Katsura, H.Nakamura and J.Thornton for helpful discussion, Drs N.Takahashi, S.Goshima, K.Kyozuka, S.Amemiya, Professor Y.Yazawa, A.Fuji and I.Yasumasu for the supply of biological materials and Misses N.Harada, Y.Ishikawa, Y.Horio, Mrs Y.Yamashita and Mrs Y.Suzuki for preparing manuscripts. I especially thank Drs A.Klug and M.Churchill for reading manuscripts and critical comments. I thank Applied Biosystem Japan for the gift of an octapeptide. I thank the following marine laboratories for preparing biological materials: Marine Biomedical Institute, Sapporo Medical College (Rishiri), Marine Biological Station, Hokkaido University (Akkeshi), Marine Biological Station, Tohoku University (Asamushi) and Marine Biological Station, University of Tokyo (Misaki). I thank EMBO for its short-term fellowship which gave me a chance to discuss my research with British scientists. This work was supported by Grants-in-Aid for Special Project Research (62105005, 62220027), Grant-in-Aid for Developmental Scientific Research (63870005) and Grant-in-Aid for Encouragement of Young Scientists (63780339) from the Ministry of Education, Science and Culture of Japan.

References

Alber,T., Dao-pin,S., Wilson,K., Wozniak,J.A., Cook,S.P. and Matthews,B.W. (1987) *Nature*, **330**, 41–46.

Baker,E.N. and Hubbard,R.E. (1984) *Prog. Biophys. Mol. Biol.*, **44**, 97–179.

Bontemps,J., Houssier,C. and Fredericq,E. (1975) *Nucleic Acids Res.*, **2**, 971–984.

Bradbury,E.M., Maclean,N. and Matthews,H.R. (1981) In *DNA, Chromatin and Chromosomes*. Blackwell Scientific Publications, Oxford.

Brayer,G.D. and McPherson,A. (1983) *J. Mol. Biol.*, **169**, 565–596.

Burlingame,R.W., Love,W.E., Wang,B.-C., Hamlin,R., Xuong,N.-H. and Moudrianakis,E.N. (1985) *Science*, **228**, 546–553.

Busslinger,M. and Barberis,A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5676–5680.

Cheng,Y.-H., Yang,J.-T. and Martinez,H.M. (1972) *Biochemistry*, **11**, 4120–4131.

Chou,P.Y. and Fasman,G.D. (1974) *Biochemistry*, **13**, 222–245.

Feughelman,M., Langridge,R., Seeds,W.E., Stokes,A.R., Wilson,H.R., Hooper,C.W., Wilkins,M.H.A., Barclay,R.K. and Hamilton,L.Q. (1955) *Nature*, **175**, 834–838.

Finch,J.T., Lewit-Bentley,A., Bentley,G.A., Roth,M. and Timmins,P.A. (1980) *Phil. Trans. R. Soc. Lon.*, Ser. B, **290**, 635–638.

Godowski,P.J., Riger,S.R., Miesfeld,R. and Yamamoto,K. (1987) *Nature*, **325**, 365–368.

Green,S. and Chambon,P. (1987) *Nature*, **325**, 75–78.

Green,S., Walter,P., Kumar,V., Krust,A., Borhret,J.-M., Argos,P. and Chambon,P. (1986) *Nature*, **320**, 134–139.

Hartman,P.G., Chapman,G.E., Moss,T. and Bradbury,M. (1977) *Eur. J. Biochem.*, **77**, 45–57.

Hartshorne,T.A., Blumberg,H. and Young,E.T. (1986) *Nature*, **320**, 283–287.

Klug,A., Rhodes,D., Smith,J., Finch,J.T. and Thomas,J.O. (1980) *Nature*, **287**, 509–516.

Kopka,M.L., Fratini,A.V., Drew,H.R. and Dickerson,R.E. (1983) *J. Mol. Biol.*, **163**, 129–146.

Kopka,M.L., Pjura,P.E., Goodsell,D.S. and Dickerson,R.E. (1987) In Eckstein,F. and Lilley,D.M.J. (eds), *Nucleic Acids and Molecular Biology*. Springer-Verlag, Berlin.
Lai,Z.-C. and Childs,G. (1986) *Nucleic Acids Res.*, **14**, 6845–6856.
Laughon,A. and Scott,M.P. (1984) *Nature*, **310**, 25–31.
Lewis,P.N., Momany,F.A. and Scheraga,H.A. (1973) *Biochem. Biophys. Acta*, **303**, 211–229.
McGhee,J.D., Nickol,J.M., Felsenfeld,G.F. and Rau,D.C. (1983) *Cell*, **33**, 831–841.
McKay,D.B. and Steitz,T.A. (1981) *Nature*, **290**, 744–749.
Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.
Poccia,D. (1987) In Schlegel,R.A. and Halleck,M.S. (eds), *Molecular Regulation of Nuclear Events in Mitosis and Meiosis*. Academic Press, Orlando, FL pp. 149–178.
Rees,D.C., Lewis,M. and Lipscomb,W.N. (1983) *J. Mol. Biol.*, **168**, 367–387.
Rhodes,D. (1985) *EMBO J.*, **4**, 3473–3482.
Richmond,T.J., Finch,J.T., Rushton,B., Rhodes,D. and Klug,A. (1984) *Nature*, **311**, 532–537.
Rosenberg,U.B., Schröder,C., Preiss,A., Kienlin,A., Coté,S., Riede,I. and Jäckle,H. (1986) *Nature*, **314**, 336–339.
Sap,J., Munoz,A., Damm,K., Goldberg,Y., Gynsdael,J., Leutz,A., Beug,H. and Vennstrom,B. (1986) *Nature*, **324**, 635–640.
Suzuki,M. (1988a) *Proc. Japan Acad.*, Ser. B, **64**, 29–32.
Suzuki,M. (1988b) *Proc. Japan Acad.*, Ser. B, **64**, 33–36.
Suzuki,M. (1989) *J. Mol. Biol.*, in press.
Takeda,Y., Ohlendorf,D.H., Anderson,W.F. and Matthews,B.W. (1983) *Science*, **221**, 1020–1026.
Tanaka,I., Appelt,K., Dijk,J., White,S.W. and Wilson,K.S. (1984) *Nature*, **310**, 376–381.
Tullius,T.D., Dombroski,B.A., Churchill,M.E.A. and Kam,L. (1987) *Methods Enzymol.*, **155**, 537–558.
von Holt,C., de Groot,P., Schwager,S. and Brandt,W.F. (1984) In Stein,G.S., Stein,J.L. and Marzluff,W.F. (eds), *Histone Genes*. Wiley-Interscience, New York, pp. 65–105.
Wilmot,C. and Thornton,J. (1989) *J. Mol. Biol.*, in press.
Zalenskaya,I.A., Pospelov,V.A., Zalensky,A.O. and Vorobev,V.I. (1981) *Nucleic Acids Res.*, **9**, 473–487.

Received on September 28, 1988; revised on December 20, 1988