J Mol Model (2005) 12: 76–84
DOI 10.1007/s00894-005-0008-8

ORIGINAL PAPER

Claudio Greco · Elena Sacco · Marco Vanoni
Luca De Gioia

Identification and in silico analysis of a new group of double-histone fold-containing proteins

Received: 25 January 2005 / Accepted: 27 June 2005 / Published online: 25 October 2005

Abstract The double-histone fold is a rare protein fold in which two consecutive regions characterized by the typical structure of histones assemble together, thus giving a histone pseudodimer. Previously, this fold was found in a few prokaryotic histones and in the regulatory region of guanine–nucleotide exchange factors of the Sos family. Standard methods of sequence comparison did not allow us to find new proteins containing a histone pseudodimer, as previously reported (Sondermann et al. 2003). However, a deeper investigation of protein sequences showed that the two histone folds included in Sos proteins share significant sequence similarity with nucleosomal histones. On the basis of this observation, we applied a specific strategy of sequence-homology search, which led to the identification of a new group of histone pseudodimers in Cca3 and proteins similar to Cca3 (Cca3S). A homology model of the histone pseudodimer included in rat Cca3 was constructed. A subsequent structure–function relationship study revealed that the histone pseudodimers included in Cca3 and Cca3S proteins, but not those present in Sos proteins, could retain the ability of mediating protein–DNA interactions, and could consequently act as DNA-binding modules.

Keywords Double histone fold · Cca3 · Sequence alignment · Homology modeling · Molecular electrostatic potential

Abbreviations Cca3: Confluent · 3Y1: Cell-associated · hSos1: Human Son of Sevenless 1 · Hs: *Homo sapiens* · Mm: *Mus musculus* · Rn: *Rattus norvegicus* · Tn: *Tetraodon nigroviridis* · Fr: *Fugu rubripes* · Xt: *Xenopus tropicalis* · Xl: *Xenopus laevis* · Dm: *Drosophila melanogaster* · Ag: *Anopheles gambiae* · Ce: *Caenorhabditis elegans* · Nr: Non-redundant

C. Greco · E. Sacco · M. Vanoni (✉) · L. De Gioia (✉)
Dipartimento di Biotecnologie e Bioscienze,
Università degli Studi Milano-Bicocca,
P.zza della Scienza 2, 20126 Milano, Italy
E-mail: marco.vanoni@unimib.it
Tel.: +39-02-64483525
E-mail: luca.degioia@unimib.it
Tel.: +39-02-64483463

Introduction

Histone folds are structural elements that are able to form dimers by means of tight interactions between hydrophobic surfaces. Normally, a histone fold is composed of a long α-helix flanked by two or three shorter helices. In the nucleosome core particle, two pairs of H2a–H2b and H3–H4 histone heterodimers assemble together, giving rise to a disk-like octamer upon which DNA rolls up [1].

The publication of the X-ray structure of the prokaryotic histone from *Methanopyrus kandleri* (PDB code: 1f1e) highlighted a novel protein fold, which is formed by the assembly of two consecutive histone folds included in the same peptide chain [2]. More recently, the publication of the X-ray structure of the amino-terminal domain of hSos1 (PDB code: 1q9c) showed that this protein module also assumes a similar fold, dubbed the histone pseudodimer [3] and here also referred to as "double-histone fold". In fact, the evolutionary relationship between the H2a histone and the domain spanning the protein sequence 96–190 of hSos1 had been already disclosed because of the high sequence similarity [4] between the two domains. However, the first histone-like domain of hSos1, spanning the protein portion 6–95, does not show any evident sequence similarity with histones [3]. Moreover, it is unknown whether the "double-histone fold" can be found in other protein families.

Prompted by these observations, we initially searched for possible remote homologues of the whole-double histone fold of hSos1. It turned out that the protein portion 22–95 shows similarity to proteins characterized by the presence of the H2b histone domain. On the basis of these results, we searched for other proteins

characterized by a sequence compatible with the "double histone fold". This search allowed us to detect histone pseudodimers in Cca3 and similar to Cca3 (Cca3S) proteins. The homology modeling and the in silico characterization of these histone pseudodimers is reported here.

Methods

The initial search for proteins containing putative histone pseudodimers was carried out with Psi-Blast [5].

Multiple sequence alignments were carried out with ClustalW, [6] and manually refined. The neighbor joining tree of protein sequences was generated using the alignment obtained with ClustalW.

Secondary-structure predictions were obtained using three different tools: Psi-Pred, [7] J-pred [8] and PHD [9]. Meta-predictions were carried out by comparing the results obtained from these three servers, and taking into account only the sequence regions that were predicted to assume a particular secondary structure by at least two servers, with a grade of reliability of 50% or higher.

Fold recognition results were obtained using the 3D-jury meta server [10].

The Moe program suite [11] was used to obtain a 3D-model of the histone pseudodimer included in rat Cca3. The H2a–H2b histone dimer from *Gallus gallus* was chosen as a template (PDB code: 1eqz). The homology-modeling procedure was based on the alignment reported in Fig. 1. The construction of the model started with the production of ten intermediate homology models, which were the result of the permutational selection of different loop candidates and side-chain rotamers. The intermediate models were averaged in order to build the final model; to this end, the Cartesian Average Method was employed. Both the intermediate models and the final model were refined by molecular mechanics. All structures were energy minimized using the Amber94 force field.

The reliability of the final model was evaluated with PROCHECK [12].

The analysis of the model was carried out with Pymol [13] and Swiss PDB viewer [14]. Swiss PDB viewer was also used to obtain the electrostatic potential maps shown in Fig. 2.

Results and discussion

Individuation of remote homology between hSos1 and histone H2b

The alignment between the domain spanning the hSos1 protein portion 22–95 and the histone H2b was analyzed to evaluate possible remote homology (Fig. 3). Remarkably, it turned out that the H2b domain and the corresponding hSos1 domain show a significant similarity (22.1% identity). The analysis of the alignment highlights a strict conservation of a group of hydrophobic residues; the amino acids included in this hydrophobic pattern are mainly localized at the relative positions i and $i+3$ or $i+4$ along each α-helix, as shown in Fig. 3. The conservation of this feature maintains the amphiphilic character of the α-helices, a crucial property for the formation of tight interactions between the two histone folds and for the correct folding of this domain [3].

Three gaps are evident in the hSos1 sequence after the alignment. The first two gaps fall in a region that corresponds to an α-helix in the H2b domain. Since the gaps correspond to four and three residues in H2b, respectively, they do not affect the proper alignment of the crucial hydrophobic residues, strongly suggesting that two helix turns are missing in the second α helix of the H2b domain of hSos1.

Identification and in silico characterization of a new group of histone pseudodimer containing proteins

Prompted by the above observations, we searched for other proteins that could be characterized by the peculiar double-histone fold observed in hSos1. Submission of the N-terminal protein sequence of hSos1 (residues 1–190) to Psi-Blast did not highlight proteins sharing a significant sequence similarity with the probe sequence, in agreement with previous observations [3]. However, the existence of homology between Sos proteins and histone H2b (see above) led us to use a different strategy to search for remote homologues, based on the submission to Psi-Blast of a "chimeric" sequence obtained by linking the protein regions corresponding to the H2b (residues 38–114) and H2a (residues 1–119) monomers of the human histone dimer. This search gave a statistically significant hit (E-value = $7\text{x}10^{-07}$) corresponding to the rat Cca3 protein (NCBI code: gi|19705529). In the same Psi-Blast entry, we identified proteins from *Homo sapiens*, *Mus musculus*, *Xenopus laevis*, *Tetraodon nigroviridis* and *Anopheles gambiae*, with significant sequence similarity to rat Cca3. We also identified two other proteins from rat and chicken tagged as "similar to Cca3", and their homologues from *H. sapiens*, *X. tropicalis* and *T. nigroviridis*. Further research in protein sequence databases resulted in the identification of proteins from *Fugu rubripes* and *Drosophila melanogaster*, both sharing significant sequence similarity with rat Cca3. The accession number of these protein sequences and their general features are shown in Tables 1 and 2.

In rat Cca3, an amino-terminal module encompassing the first 97 residues precedes the region of homology to histones, which is followed by ankyrin repeats and a POZ domain (Fig. 4). Ankyrin repeats and POZ domains are able to mediate protein–protein interactions in many different classes of protein [15, 16]. To compare the Cca3 protein from rat with all other putative Cca3 proteins identified in different species, we examined the pairwise

Fig. 1 **a** Multiple sequence alignment among the "chimeric" H2b–H2a histone, hSos1, and a group of Cca3 and similar to Cca3 (Cca3S) proteins. Above the alignment, are shown the positions of α-helices in hSos1 (*light blue boxes*), histone H2b (*green boxes*) and histone H2a (*yellow boxes*). Red letters indicate regions predicted to assume an α-helical structure, based on a comparison among the results obtained from three different secondary structure prediction servers (PHD, J-pred, Psi-pred; see Methods). Underlined residues belong to the pattern of aminoacids whose hydrophobicity is strictly conserved. **b** Multiple sequence alignment among histones and a group of Sos homologs. In histone H2b and H2a sequences, the residues forming hydrogen bonds with DNA are marked with *red letters*. *Red* and *light blue letters* in Sos sequences indicate conservation and conservative substitution, respectively. *Gray letters* in Sos sequences signalize regions in which there is no structural overlap between Sos and histones. Above the alignment, are shown the positions of α-helices in histone H2b (*green boxes*) and histone H2a (*yellow boxes*). **c** Multiple sequence alignment among histones and a group of Cca3 and similar to Cca3 (Cca3S) proteins. In histone H2b and H2a sequences, the residues forming hydrogen bonds with DNA are marked with red letters. *Red* and *light blue letters* in Cca3S and Cca3 protein sequences indicate conservation and conservative substitution respectively. Above the alignment, are shown the positions of α-helices in histone H2b (*green boxes*) and histone H2a (*yellow boxes*)

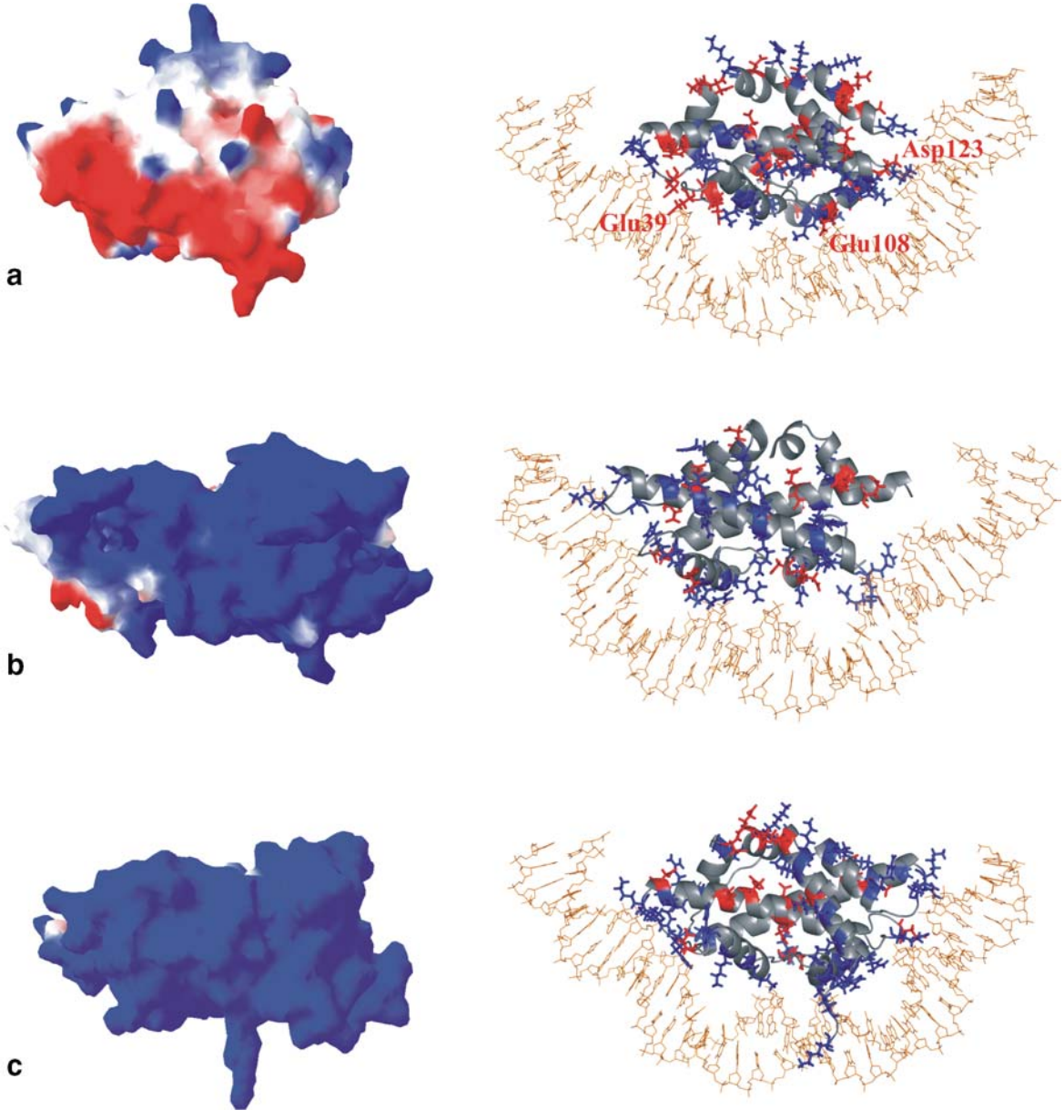

Fig. 2 **a** *Left side*: Molecular electrostatic potential of hSos1 double histone fold, in the region of putative contact between the protein and DNA. *Blue surfaces* are positively charged, while *red surfaces* are negatively charged. *Right side*: positions of basic residues (Arg, Lys, His) and acid residues (Glu, Asp) in the double histone fold of hSos1. The side chains of these residues are colored in light blue and red, respectively. The hypothetical relative positions of hSos1 (*in grey*) and DNA (*in orange*) were obtained through a superimposition of hSos1 double histone fold on the H2b–H2a nucleosomal histone dimer. **b** *Left side*: molecular electrostatic potential of rat Cca3 model, in the region of putative contact between the protein and DNA. *Right side*: positions of basic residues (Arg, Lys, His) and acid residues (Glu, Asp) in the 3D model of rat Cca3. The hypothetical relative positions of Cca3 (*in grey*) and DNA (*in orange*) were obtained through a superimposition of Cca3 double histone fold on the H2b–H2a nucleosomal histone dimer. **c** *Left side*: molecular electrostatic potential of the nucleosomal histone dimer H2b–H2a, in the region of contact between the dimer and DNA. *Right side*: positions of basic residues (Arg, Lys, His) and acid residues (Glu, Asp) in the dimer

alignment between the whole sequences of these proteins. It turns out that the amino-acid similarity is always above 50%, a threshold conventionally used to classify a group of genes as a gene family.

Then, in order to characterize the hypothetical histone pseudodimer present in the different members of the Cca3 protein family, we aligned such sequences (Fig. 1), as explained in the Methods section. In Cca3 proteins, the region similar to histone H2b was always followed by a region characterized by similarity to histone H2a. The percentages of identity among histone H2a, histone H2b and Cca3 proteins are reported in Table 1. The analysis of the alignment reveals that, considering the region similar to H2b, the amphipatic character of the putative α-helices is well-conserved, even when the sequence identity is as low as 16%.

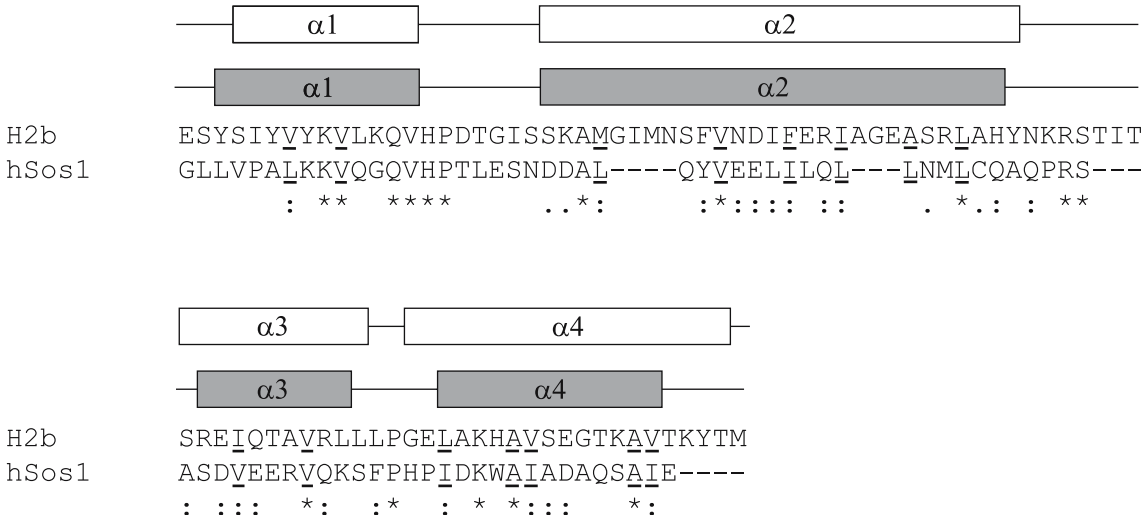

Fig. 3 Alignment between hSos1 (residues 22-95) and histone H2b (residues 38-114). *White boxes* and *grey boxes* indicate the positions of α-helices in the structure of histone H2b and hSos1, respectively. *Underlined* residues belong to the pattern of aminoacids whose hydrophobicity is strictly conserved. The alignment was obtained with ClustalW and manually refined

As for the protein region corresponding to H2a, the sequence identity with Cca3 is about 30%, and the typical pattern of hydrophobic residues is also conserved. Moreover, glycine and proline residues, which are known to be found often in loop regions, are mainly located among the regions that should correspond to α-helices.

Psi-Blast searches also showed the existence of a human homolog of Cca3 (NCBI code gi|21754934), which surprisingly lacks the region of homology to his-

Table 1 Members of the Cca3 family

Seq.N.	NCBI code	Organism	Protein (aa)	Position DHF	Percent of identity to histone H2b	Percent of identity to histone H2a
1	gi|19705529	*R. norvegicus*	1,024	98–282	27.4	26.2
2	gi|49402267[a]	*M. musculus*	1,024	98–282	27.4	26.2
3	gi|54038583[a]	*X. laevis*	1,016	94–274	24.2	29.3
4	gi|47215237[a]	*T. nigroviridis*	1,093	93–285	30.2	27.6
5	gi|31240747[a]	*A. gambiae*	1,352	386–563	16.4	37.6
6	gi|45553229[a]	*D. melanogaster*	1,326	356–527	16.4	33.8
7	SINFRUP00000078440[b]	*F. rubripes*	936	90–273	24.4	27.7
8	SINFRUP00000051335[b]	*F. rubripes*	975	95–276	20.6	27.7
9	SINFRUP00000077108[b]	*F. rubripes*	1,002	93–284	30.4	29.9

DHF Double histone fold
[a]These proteins were identified from the nr database, as uncharacterized protein sequences; accession numbers as given in the NCBI database are shown.
[b]These proteins were identified from the translated version of the genomic database of *Fugu rubripes*

Table 2 Members of the "similar to Cca3" family

Seq.N.	NCBI code	Organism	Protein (aa)	Position DHF	Percent of identity to histone H2b	Percent of identity to histone H2a
1	gi|51471006	*H. sapiens*	332	98–258	21.0	36.4
2	gi|34862936	*R. norvegicus*	316	98–268	22.1	33.8
3	gi|50784195	*G. gallus*	459	95–275	26.1	27.4
4	gi|52139159[a]	*X. tropicalis*	294	95–274	27.7	28.0
5	gi|47218320[a]	*T. nigroviridis*	323	95–277	22.4	26.7
6	gi|47225780[a]	*T. nigroviridis*	392	89–272	23.3	27.9

DHF Double histone fold
[a]These proteins were identified from the nr database, as uncharacterized protein sequences; accession numbers as given in the NCBI database are shown

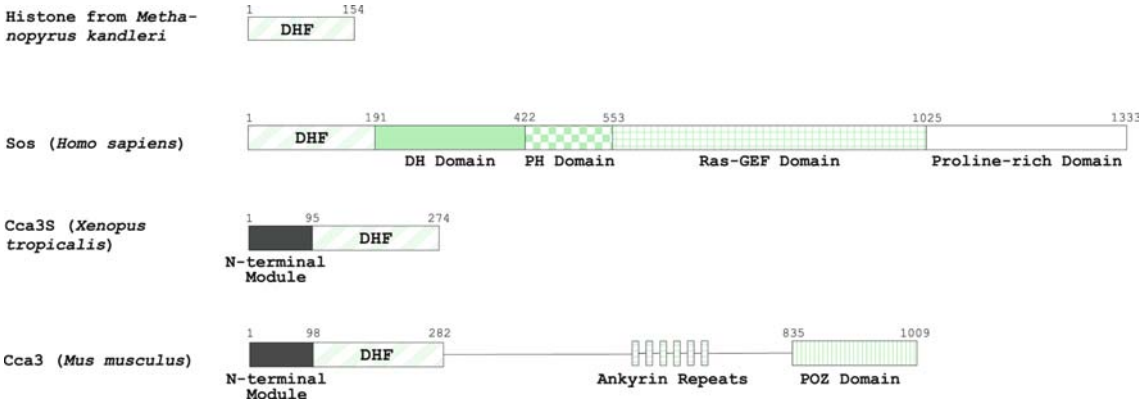

Fig. 4 Graphical representation of the four classes of proteins (archeal histone pseudodimers, Sos proteins, Similar to Cca3 and Cca3 proteins) encompassing a double histone fold (*DHF*) domain

tone H2b. Prompted by this observation, we identified the corresponding gene in the human genome sequence, and translated the nucleotidic sequence upstream of the region of homology to histone H2a. The resulting protein sequence shares 97.8% identity with the region similar to histone H2b of rat Cca3, indicating that the human Cca3 protein should also encompass two distinct histone folds.

As stated above, the Psi-Blast search found other proteins, referred to as "similar to Cca3" (Cca3S), which could be characterized by the double histone fold. These proteins are significantly shorter than Cca3 (see Table 2), and do not include ankyrin repeats or POZ domains. However, Cca3S proteins show a significant sequence similarity (above 50%) in the region corresponding to the putative double-histone fold, and in the upstream amino-terminal module. Furthermore, Cca3S proteins show the expected conservation of the amphipatic character of the putative α-helices.

The alignment among Cca3S proteins and histones showed the absence of gaps in the putative α-helices, apart from the case of rat Cca3S which lacks 14 residues in the region of homology to histone H2a. This gap could cause the fusion of two helices, namely the third, longest helix of this region and the following one. This fusion and the consequent loss of an α-helix flanking the longest helix in one of the two histone folds is expected to be compatible with the formation of the double-histone fold, since some transcription factors are composed of couples of interacting histone folds, in which the first monomer presents a complete histone fold (two short helices flanking a longer one), while the second monomer lacks one of the shorter helices (see [17] and 1h3o PDB record, shown in Fig. 5). Similar structural features are also expected to characterize the human version of Cca3S, in which the H2a-like histone fold comprises only two putative α-helices (see Fig. 1).

All these observations strongly suggest a common fold for all these domains. To investigate this aspect further, we predicted the secondary structure for these domains, as explained in the Methods section, and compared it with experimental data for hSos1 and the H2a and H2b proteins of the human histone dimer. The comparison among the results obtained from three different prediction servers gave rise to a meta-prediction which supports the hypothesis that Cca3 and Cca3S proteins assume a secondary structure very similar to that of hSos1 and histones H2a and H2b (Fig. 1).

To test the previous observations further, we also submitted to the fold recognition meta-server 3D-Jury (see Methods) the protein portions of Cca3 and Cca3S which should have structural features similar to the double histone fold domain of hSos1. The results of this analysis confirmed that Cca3 and Cca3S protein sequences are compatible with the structure of the amino-terminal domain of hSos1 (data not shown).

Analysis of the structure–function relationship in Sos, Cca3, and Cca3S histone pseudodimers

The presence of double-histone fold domains in Sos, Cca3, and Cca3S proteins rises an important question: are these structural modules implicated in DNA binding? To investigate this possibility, we carried out a deep in silico analysis on Sos, Cca3, and Cca3S histone pseudodimers. The agreement between the results of our computational investigation and the experimental data available in the literature was then checked.

Sos proteins are ubiquitously expressed Ras–GEFs (guanine nucleotide exchange factors). The activation of growth factor receptors stimulates the translocation of Sos proteins from the cytoplasm to the plasma membrane, where Ras is localized. The interaction between Sos and Ras at the membrane causes Ras to release GDP, bind GTP and become activated [18]. Since there is no evidence supporting the possibility that Sos proteins can move to the nucleus, the histone domain is probably not involved in DNA binding. In fact, our analysis based on sequence comparisons among histones and the different Sos orthologs showed a substantial lack of conservation in the

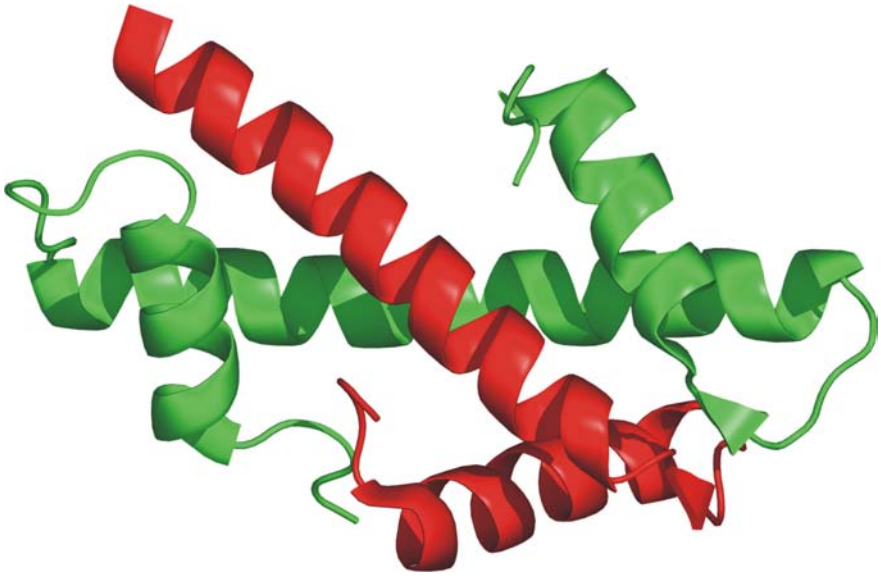

Fig. 5 1h3o PDB record; the histone fold that lacks one of the short helices flanking the longest helix is colored in red

pattern of residues that would be directly involved in the interaction with DNA (Fig. 1). In the H2a–H2b nucleosomal histone dimer, the surface region that mediates protein-DNA contacts is strictly basic. As a result, attractive interactions between the histone dimer and deoxyribonucleic acids can take place. The corresponding surface region of Sos histone pseudodimer proved to be prevalently negatively charged, as shown in the electrostatic potential map shown in Fig. 2. Moreover, in the hypothetical complex between hSos1 and DNA, three acid residues of hSos1 (Glu39, Glu108, Asp123) would be in direct contact with phosphate groups of DNA (Fig. 2).

Analogous computational analyses were carried out on the Cca3 and Cca3S proteins. The alignment among the sequences of all these proteins and histones highlights the fact that the residues that could mediate the interaction with DNA are prevalently conserved or substituted with other amino acids that could be involved in DNA binding. In particular, a cluster of conserved basic residues is observed in the amino-terminal part of the region of homology to histone H2a (Fig. 1). In the case of Cca3 proteins, the presence of a POZ domain also suggests a tight relationship between the double-histone fold domain and DNA, since POZ domains are commonly found in nuclear proteins containing other well-established DNA-binding modules, belonging to the zinc-finger domain family [19].

In order to shed light on some key structural features, we used a homology-modeling procedure on the double-histone fold of rat Cca3 (see Methods), using the H2b-H2a histone dimer as a template. The structural

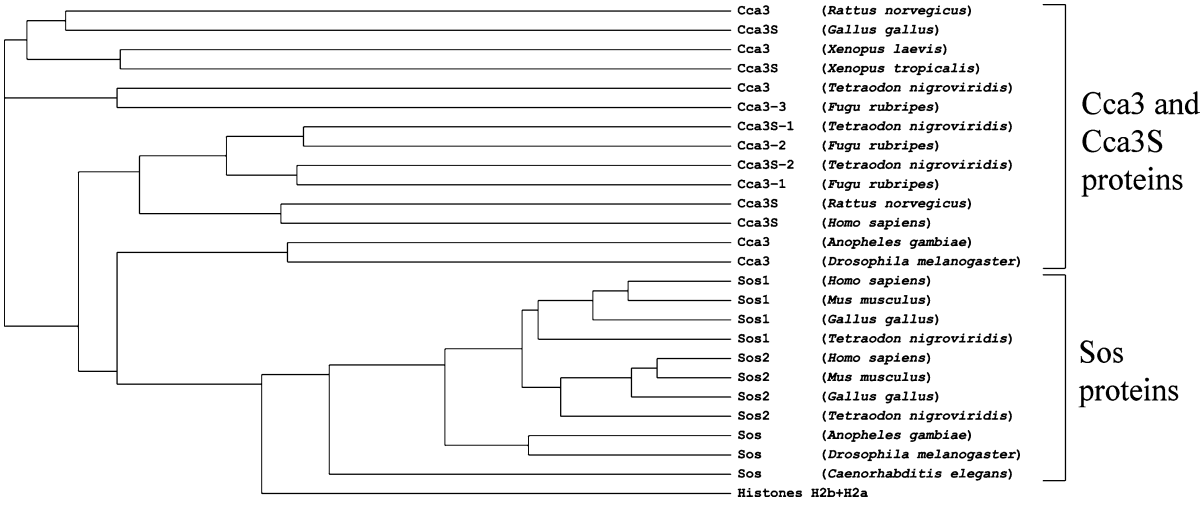

Fig. 6 Organization of histones H2b and H2a, Sos, Cca3 and Cca3S proteins based on molecular phylogenetics. The tree is generated using the entire double histone fold domain of Sos, Cca3 and Cca3S proteins, together with the chimeric sequence obtained linking the protein regions corresponding to the H2b (residues 38–114) and H2a (residues 1–119) monomers of the human histone dimer

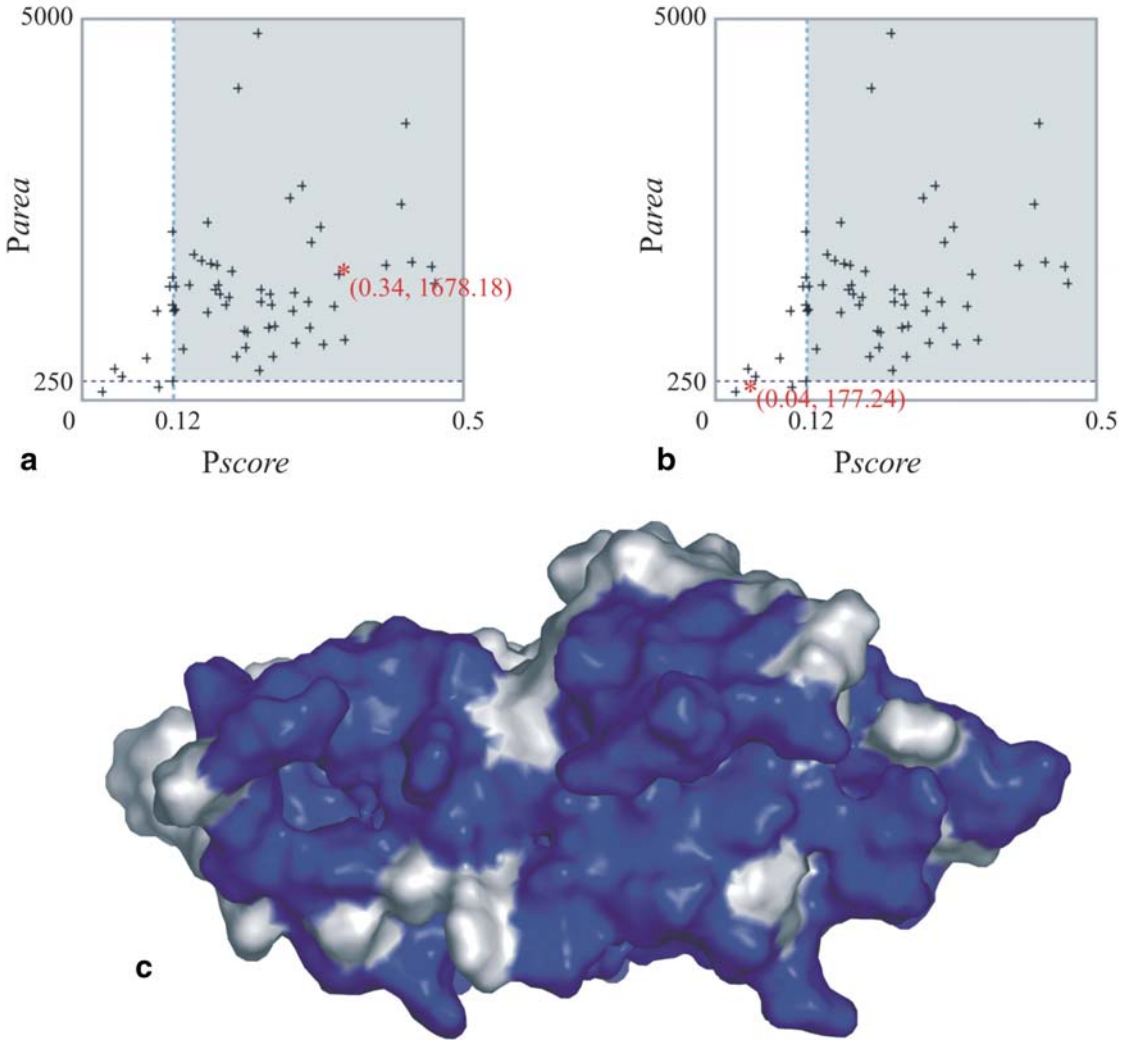

Fig. 7 **a** and **b** Graphical representation of the statistical parameters (Pscore and Parea, see [20]) on which the prediction of DNA-binding site is based. Black crosses indicate the Pscore and Parea values calculated for 63 representative dsDNA-binding proteins, while the red asterisks refer to the values of the same parameters for Cca3 histone pseudodimer model **(a)**, and for the amino-terminal domain of hSos1 **(b)**. Only the proteins with Pscore > 0.12 and Parea > 250 (thus included in the upper right region of the graph) are considered dsDNA-binding proteins. **c** Localization of the predicted DNA-binding surface (*in blue*) on the rat Cca3 model

reliability of the model was evaluated by means of PROCHECK; the calculation of PROC-AVE parameter, (which represents a carefully weighted average of all the analyses performed by PROCHECK) gave a value of −0.27, significantly higher than the threshold of −0.5 which divides poor and good models. The rat Cca3 model shows that the putative region of interaction between the double-histone fold and DNA is very rich in basic residues and no acid residues in direct contact with DNA are found (Fig. 2). Moreover, the calculation of the electrostatic potential map of the model highlights that the putative region involved in DNA binding is positively charged (Fig. 2). This conservation of physico-chemical properties in the putative region of contact to DNA is even more significative considering the fact that the Cca3 and Cca3S proteins are more distantly related to histones than the Sos proteins (Fig. 6).

The availability of a model for the double-histone fold included in rat Cca3 allowed us to apply a novel and highly reliable computational method for the identification of DNA-binding proteins. This method, developed by Tsukiya et al. [20] focuses on the shape of the molecular surface of the protein and DNA and on the electrostatic potential on the surface. The resulting prediction scheme shows 86 and 96% accuracy for DNA-binding and non-DNA-binding proteins, respectively [20]. It is important to note that this method is complementary with respect to the phylogenetic and comparative approach applied in the present work up to now. The results obtained from the application of such a method confirmed all the observations reported above: the histone pseudodimer included in hSos1 was indicated as a non-DNA-binding domain, while the double histone fold from rat Cca3 was recognized as a

DNA-binding module (Fig. 7). Furthermore, the surface portion indicated by the algorithm as the DNA-binding region on Cca3 histone pseudodimer lies over the conserved basic surface described above (see Figs. 7, 2).

The above observations strongly support the hypothesis that the double-histone fold of the Cca3 and Cca3S proteins could mediate the interaction between these proteins and deoxyribonucleic acids. Cca3-encoding mRNA preferentially accumulates in growth-arrested 3Y1 line cells, [21] and its level decreases transiently before initiation of DNA replication in regenerating rat-liver cells [22]. No molecular information is available so far regarding the molecular details of the function of Cca3 proteins. In light of our computational data, it is tempting to speculate that the pattern of expression of Cca3 may correlate with preferential binding of Cca3 to DNA in a pre-replicative condensed G0/G1 state. Thus, biochemical studies of proteins of the Cca3 family will receive novel impulse and directions from our studies and will contribute to shed new light on the function of this interesting class of proteins.

Conclusions

In the double-histone fold, two consecutive regions characterized by the typical structure of histones assemble together, thanks to tight interactions between their hydrophobic surfaces [3]. Until now, this protein fold had been identified only in prokaryotic histones and in Sos proteins. Sequence comparisons, secondary structure predictions and fold recognition methods indicate that this protein fold is more widespread than previously expected. In fact, histone pseudodimers have been identified in conserved sequence modules of Cca3 and "similar to Cca3" proteins.

Previous literature data indicated that the prokaryotic histone pseudodimer could have a role in chromatin packaging, [23] while the double-histone fold domain in Sos proteins is known to be implicated in the mediation of intermolecular interactions that lead to the autoinhibition of the Ras-GEF activity [24]. In silico analysis of Cca3 and Cca3S has allowed us to show that the physicochemical properties of the histone pseudodimers in these proteins are compatible with DNA binding. This result suggests that the double-histone fold could also be involved in mediating the interactions between DNA and multidomain proteins. Moreover, our computational data will be useful for planning further experimental functional characterization of these novel classes of proteins.

The analysis of new genomic sequences and the availability of more sensitive computational tools will allow us to verify if other protein families can be included in the group of histone-pseudodimer containing proteins.

Acknowledgements M.V. thanks MIUR-FAR for financial support. E.S. was supported by a FIRC fellowship.

References

1. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Nature 389:251–260
2. Fahrner RL, Cascio D, Lake JA, Slesarev A (2001) Protein Sci 10:2002–2007
3. Sondermann H, Soisson SM, Bar-Sagi D, Kuriyan J (2003) Structure 11:1583–1593
4. Baxevanis AD, Arents G, Moudrianakis EN, Landsman D (1995) Nucleic Acids Res 23:2685–2691
5. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Nucleic Acids Res 25:3389–3402
6. Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ (1994) Nucleic Acids Res 22:4673–4680
7. Jones DT (1992) J Mol Biol 292:195–202
8. Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ (1998) Bioinformatics 14:892–893
9. Rost B, Sander C, Schneider R (1994) Comput Appl Biosci 1:53–60
10. Ginalski K, Elofsson A, Fischer D, Rychlewski L (2003) Bioinformatics 19:1015–1018
11. http://www.chemcomp.com/
12. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) J Appl Cryst 26:283–291
13. DeLano LW, DeLano Scientific LLC, San Carlos, CA, USA
14. Guex N, Peitsch MC (1997) Electrophoresis 18:2714–2723
15. Bork P (1993) Proteins 17:363–374
16. Ahmad KF, Engel CK, Prive GG (1998) Proc Natl Acad Sci USA 95:12123–12128
17. Werten S, Mitschler A, Romier C, Gangloff YG, Thuault S, Davidson I, Moras D (2002) J Biol Chem 277:45502–45509
18. Schlessinger J (2000) Cell 103:211–215
19. Collins T, Stone JR, Williams AJ (2001) Mol Cell Biol 21:3609–3615
20. Tsuchiya Y, Kinoshita K, Nakamura H (2004) Proteins 55:885–894
21. Hayashi Y, Kiyono T, Fujita M, Ishibashi M (1997) Biochim Biophys Acta 1352:145–50
22. Hayashi Y, Ichinose M, Yuasa H, Tatematsu M, Ishibashi M (1997) FEBS Lett 406:147–150
23. Slesarev AI, Belova GI, Kozyavkin SA, Lake JA (1998) Nucleic Acids Res 26:427–430
24. Jorge R, Zarich N, Oliva JL, Azañedo M, Martínez N, de la Cruz X, Rojas JM (2002) J Biol Chem 277:44171–44179