Hypothesis

How do histone acetyltransferases select lysine residues in core histones?

Akatsuki Kimura[a], Masami Horikoshi[a,b,*]

[a] *Laboratory of Developmental Biology, Department of Cellular Biology, Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan*

[b] *Horikoshi Gene Selector Project, Exploratory Research for Advanced Technology, Japan Science and Technology Corporation, 5-9-6 Tokodai, Tsukuba, Ibaraki, Japan*

Received 5 May 1998; revised version received 8 June 1998

Abstract Acetylation of specific lysines of core histone N-terminal tails correlates with chromatin assembly and specific regulation of gene expression. Core histones acetylated at particular lysines mediate effects on chromatin function; however, the manner in which different histone acetyltransferases (HATs) discriminate lysines is unknown. Here we propose a putative rule for lysine selection by HATs based on the primary sequence in the vicinity of lysines in the core histone N-terminal tails and in flanking sequence. This provides insight into the molecular basis of site selection of core histones by HATs.

Key words: Histone acetyltransferase; Histone; Site specificity; Chromatin

1. Introduction

Since the early discovery of histone acetylation [1], histone acetyltransferases (HATs), enzymes responsible for this modification, have been considered modifiers of chromatin structure and key players of gene expression [2]. Recent identification of various types of HATs [3–12] represents a significant breakthrough in understanding the functional role of acetylation. Fifteen lysines (in human cells) which reside in four core histones are acetylated in various patterns in vivo [13,14]. Site specificity of different HATs is also divergent [6,8,11,15,16]. Taking into consideration that acetylation of particular lysines in core histones mediates specific effects on chromatin function [17], it is important to elucidate the molecular basis involved in selection of specific lysines by HATs. The relationship between the divergent pattern of lysine acetylation of core histones and site specificity of HATs has remained obscure.

Our analysis of the primary sequence in the vicinity of lysines of the core histone N-terminal tails showed that lysines acetylated in vivo can be classified into six groups and this classification fits well with site specificity of known HATs and patterns of acetylation in vivo.

2. Lysines acetylated in vivo can be classified into six groups

Two criteria were introduced to classify the 15 lysines acetylated in vivo in all four core histones into six groups. First, these 15 lysines were grouped into three classes on the basis of chemical properties of amino acid residues located N-terminal to the lysines (class I–III). Each was further classified into two groups according to similarities in flanking sequences (group A–F). When we compared this classification with the site specificity of HATs (Table 1), we found that lysines of a given group are acetylated by the same HATs. yHat1p acetylates only lysines of group A but not groups B–F, and yGcn5p acetylates only lysines of group B but not lysines of other groups. Site specificity of dTAF230 and SRC-1, which have been analyzed only for histone H3, also fits well with our classification. p300/CBP, general co-activators that interact with various types of DNA binding regulatory factors, acetylate lysines of more than two groups implying that p300/CBP are HATs with a lesser site specificity. Since the site specificity of HATs analyzed to date correlates well with our classification, we propose that distinct HATs have the potential to distinguish among groups A–F.

Examination of all 28 lysines residing in the core histone N-terminal tails revealed that while lysines acetylated in vivo can be classified into groups A–F, lysines that are not acetylated in vivo cannot be classified into these groups, based on our criteria, except for Lys^{9} of human histone H2A (H2A-9), which is classified into group B (Table 2). H2A-9 was found to be acetylated in calf thymus [18], hence, HATs may also recognize this residue, under certain conditions. The finding that lysines not acetylated in vivo cannot be classified into groups A–F supports our proposal that classification can be applied to all HATs active in vivo.

3. Discussion

Since our classification fits well with site specificity of known HATs and also with patterns of acetylation in vivo, residues critical for our classification may also be important for other unknown HATs to distinguish among lysines of the four core histones. Therefore, we predict the existence of novel HATs responsible for acetylation of specific lysines which belong to groups C, D, E and/or F.

Mechanisms involved in site specificity of HATs in complex [15,19] will require further study. However, taking into consideration that yeast Hat1p and Hat2p complex acetylates H4-12 (sequence KGLG<u>K</u>), but not H2A-5 and H4-5 (SGRG<u>K</u>) [15], the groups may be further classified into subgroups according to differences in properties of the other residues in the flanking region of these acetylated lysines. However, there are other proteins which have been reported to be the targets of HATs that contain sequences which do not fall into the proposed groups [20,21]. Thus, there remains the possibility of other ways of substrate recognition by HATs.

Although recognition of substrate by enzyme is determined

*Corresponding author. Fax: (81) (3) 5684 8341.
E-mail: horikosh@imcbns.iam.u-tokyo.ac.jp

PII: S0014-5793(98)00752-2

Table 1
Comparison among classified groups of lysines and site specificity of HATs

class[b]	group	position	flanking sequence[c]	site specificity of HATs[a] yHat1p [15]	yGcn5p [16]	dTAF230 [8]	SRC-1 [11]	p300/CBP [6]
I (G/A)	A	H2A- 5	SGRGKQGGK	?	-	-	-	?
		H4 - 5	SGRGKGGKG	+	-	?	?	+
		H4 -12	KGLGKGGAK	+	-	?	?	+
	B	H3 -14	STGGKAPRK	-	+	+	+	?
		H4 - 8	GKGGKGLGK	-	+	?	?	+
		H4 -16	KGGAKRHRK	-	+	?	?	+
II (S/T)	C	H2B- 5	PEPSKSAPA	-	-	-	-	?
		H2B-15	KKGSKKAIT	-	-	-	-	?
	D	H2B-20	KAITKAQKK	-	-	-	-	?
		H3 - 4	ARTKQTAR	-	-	-	-	?
		H3 -23	QLATKAARK	-	-	?	?	?
III (K/R)	E	H2B-12	PAPKKGSKK	-	-	-	-	?
		H3 -18	KAPRKQLAT	-	-	-	-	?
	F	H3 - 9	QTARKSTGG	-	-	?	?	?
		H3 -27	KAARKSAPA	-	-	?	?	?

The amino acid sequences of human histones are shown [13,14]. The site specificity of HATs are analyzed for free histones or amino-terminal tail peptides.
[a]Abbreviations: +, acetylated; −, not acetylated; ?, not determined.
[b]Amino acid residues N-terminal to the acetylated lysines are parenthesized.
[c]Acetylated lysines are underlined, residues similar within each group are in bold.

Table 2
Classification of lysines residing in the core histone N-terminal tails

	class I	class II	class III	others
acetylated *in vivo*	A SGRGKQGGK (H2A-5) SGRGKGGKG (H4-5) KGLGKGGAK (H4-12) B STGGKAPRK (H3-14) GKGGKGLGK (H4-8) KGGAKRHRK (H4-16)	C PEPSKSAPA (H2B-5) KKGSKKAIT (H2B-15) D KAITKAQKK (H2B-20) ARTKQTAR (H3-4) QLATKAARK (H3-23)	E PAPKKGSKK (H2B-12) KAPRKQLAT (H3-18) F QTARKSTGG (H3-9) KAARKSAPA (H3-27)	
not acetylated *in vivo*	KARAKAKTR (H2A-13) RAKAKTRSS (H2A-15) KKDGKKRKR (H2B-27) KQGGKARAK (H2A-9)[a]		KGSKKAITK (H2B-16) KAQKKDGKK (H2B-24) KDGKKRKRS (H2B-28) GKKRKRSRK (H2B-30) GGVKKPHRY (H3-37) KRHRKVLRD (H4-20)	APAPKKGSK (H2B-11) TKAQKKDGK (H2B-23) TGGVKKPHR (H3-36)

The amino acid sequences of human histones are shown [13,14]. Lysines examined are underlined and their positions are parenthesized, residues similar within each group/class are in bold.
[a]This lysine can be classified into group B (discussed in the text).

by the tertiary structure of involved protein, our classification of lysines as based on flanking sequence provides insight into the manner in which HATs select particular lysines of core histones, which may possibly be applied to substrate selection by histone deacetylases (HDACs).

Acknowledgements: We thank Dr. T. Suzuki and M. Ohara for helpful comments.

References

[1] Allfray, V.G., Faulkner, R. and Mirsky, A.E. (1964) Proc. Natl. Acad. Sci. USA 51, 786–794.
[2] Wade, P.A., Pruss, D. and Wolffe, A.P. (1997) Trends Biochem. Sci. 22, 128–132.
[3] Kleff, S., Andrulis, E.D., Anderson, C.W. and Sternglanz, R. (1995) J. Biol. Chem. 270, 24674–24677.
[4] Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996) Cell 84, 843–851.
[5] Yang, X.-J., Ogryzko, V.V., Nishikawa, J., Howard, B.H. and Nakatani, Y. (1996) Nature 382, 319–324.
[6] Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. and Nakatani, Y. (1996) Cell 87, 953–959.
[7] Bannister, A.J. and Kouzarides, T. (1996) Nature 384, 641–643.
[8] Mizzen, C.A., Yang, X.-J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y. and Allis, C.D. (1996) Cell 87, 1261–1270.
[9] Wang, L., Mizzen, C., Ying, C., Candau, R., Barlev, N., Brownell, J., Allis, C.D. and Berger, S.L. (1997) Mol. Cell. Biol. 17, 519–527.

[10] Chen, H., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y. and Evans, R.M. (1997) Cell 90, 569–580.
[11] Spencer, T.E., Jenster, G., Burcin, M.M., Allis, C.D., Zhou, J., Mizzen, C.A., McKenna, N.J., Onate, S.A., Tsai, S.Y., Tsai, M.-J. and O'Malley, B.W. (1997) Nature 389, 194–198.
[12] Yamamoto, T. and Horikoshi, M. (1997) J. Biol. Chem. 272, 30595–30598.
[13] Marvin, K.W., Yau, P. and Bradbury, E.M. (1990) J. Biol. Chem. 265, 19839–19847.
[14] Thorne, A.W., Kmiciek, D., Mitchelsom, K., Sautiere, P. and Crane-Robinson, C. (1990) Eur. J. Biochem. 193, 701–713.
[15] Parthun, M.R., Widom, J. and Gottschling, D.E. (1996) Cell 87, 85–94.
[16] Kuo, M.-H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996) Nature 383, 269–272.
[17] Turner, B.M., Birley, A.J. and Lavender, J. (1992) Cell 69, 375–384.
[18] Doenecke, D. and Gallwitz, D. (1982) Mol. Cell. Biochem. 44, 113–128.
[19] Grant, P.A., Duggan, L., Côté, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., Berger, S.L. and Workman, J.L. (1997) Genes Dev. 11, 1640–1650.
[20] Imhof, A., Yang, X.-J., Ogryzko, V.V., Nakatani, Y., Wolffe, A.P. and Ge, H. (1997) Curr. Biol. 7, 689–692.
[21] Gu, W. and Roeder, R.G. (1997) Cell 90, 595–606.