

• LIVER CANCER •

Prediction of HLA-A2-restricted CTL epitope specific to HCC by SYFPEITHI combined with polynomial method

Hai-Long Dong, Yan-Fang Sui

Hai-Long Dong, Yan-Fang Sui, Department of pathology, Fourth Military Medical University, Xi'an 710032, Shaanxi Province, China
Supported by Natural Scientific Foundation of China, No. 39830420

Correspondence to: Professor Yan-Fang Sui, Department of Pathology, Fourth Military Medical University, Xi'an 710032, Shaanxi Province, China. suiyanf@fmmu.edu.cn

Telephone: +86-29-3374541-211 **Fax:** +86-29-3374597

Received: 2003-11-04 **Accepted:** 2003-12-16

Abstract

AIM: To predict the HLA-A2-restricted CTL epitopes of tumor antigens associated with hepatocellular carcinoma (HCC).

METHODS: MAGE-1, MAGE-3, MAGE-8, P53 and AFP were selected as objective antigens in this study for the close association with HCC. The HLA-A*0201 restricted CTL epitopes of objective tumor antigens were predicted by SYFPEITHI prediction method combined with the polynomial quantitative motifs method. The threshold of polynomial scores was set to -24.

RESULTS: The SYFPEITHI prediction values of all possible nonamers of a given protein sequence were added together and the ten high-scoring peptides of each protein were chosen for further analysis in primary prediction. Thirty-five candidates of CTL epitopes (nonamers) derived from the primary prediction results were selected by analyzing with the polynomial method and compared with reported CTL epitopes.

CONCLUSION: The combination of SYFPEITHI prediction method and polynomial method can improve the prediction efficiency and accuracy. These nonamers may be useful in the design of therapeutic peptide vaccine for HCC and as immunotherapeutic strategies against HCC after identified by immunology experiment.

© 2005 The WJG Press and Elsevier Inc. All rights reserved.

Key words: Hepatocellular carcinoma; HLA-A*0201; Cytotoxic T Lymphocyte; Epitope

Dong HL, Sui YF. Prediction of HLA-A2-restricted CTL epitope specific to HCC by SYFPEITHI combined with polynomial method. *World J Gastroenterol* 2005; 11(2): 208-211
<http://www.wjgnet.com/1007-9327/11/208.asp>

INTRODUCTION

Cancer immunotherapy has become one of the most important therapeutic procedures in the past two decades. Induction of potent anti-tumor cytotoxicity T lymphocytes (CTL) responses can result in regression and prevention of metastasis, as demonstrated in experimental model systems. Thus, efforts towards the development of cancer immunotherapy have recently focused on the generation of tumor-specific T cell

responses. Over the last few years, several human genes that code for tumor antigens recognized by autologous CTL have been isolated^[1-3] and the epitopes derived from these tumor antigens have been further identified to serve as targets for CTL in the context of HLA class I molecules^[4].

Hepatocellular carcinoma (HCC) is one of the commonest malignant diseases in China and some other parts of Asia. Although encouraging advances have been made in the past few decades, there are still many difficulties in treating advanced-stage patients and in preventing recurrence and metastasis^[5]. Immunotherapy is considered promising despite the fact that it is still a supplementary method in clinical practice^[6,7]. The frequency of HLA-A2 allele in HCC patients is 53.5% in Chinese population^[8]. It is very valuable to identify the tumor antigen epitopes which are presented by HLA-A2 and able to induce epitope-specific CTL against HCC cells.

In this study, we report a simple and efficient prediction method to identify candidate HLA-A2 restricted CTL epitopes from the tumor antigen that is closely associated with HCC.

MATERIALS AND METHODS

Materials

MAGE-1 (309aa), MAGE-3 (314aa), MAGE-8 (234aa), P53 (393aa) and AFP (609aa) were selected as objective antigens in this study for the close association with HCC. The amino acid sequences of these tumor antigens were quoted from EMBL database.

Methods

SYFPEITHI prediction Database retrieval can be performed on any HTML-browser supporting JavaScript. The main page of the database (<http://www.uni-tuebingen.de/uni/kxi/>) offers three sections: "Find Your Motif", "Epitope prediction" and "Information". After a preselection of one or multiple MHC-types, the "Epitope prediction" section allows the user to predict candidate epitope from protein sequence of a tumor antigen. The sequences of the protein or its gene, the restriction element are available.

The frequent of HLA-A2 types in HCC patients is 53.5% in the Chinese population and HLA-A*0201 is the major subtype of HLA-A2 allele, the HLA-A*0201 type was chosen from the frame of "Select MHC type". In this study, the nonamers (9aa) were chosen from the frame of "Choose a nonamer" due to the typical length of a class I ligand comprised 9 amino acids (nonamers). Following the amino acid sequences of each tumor antigen were inputted, the epitope prediction program was processed immediately. The values of all possible nonamers of a given sequence were added together and the ten high-scoring peptides of each antigen were selected as optimal T-cell epitope for further study.

Polynomial method analysis

The basic premise of this method is independent binding of individual side-chains (IBS). When residue *R* occurs at position *i* in the peptide, it is assumed to contribute a constant amount *R_i* to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. Parameters *R_i* are estimated

Table 1 Coefficients for the polynomial method

Res	Position								
	1	2	3	4	5	6	7	8	9
A	-2.38	-3.22	-2.8	-2.66	-2.89	-2.7	-2.35	-3.07	-2.49
C	-2.94	-15.0	-2.58	-1.96	-3.29	-2.22	-2.97	-2.37	-15.0
D	-3.69	-15.0	-3.46	-2.71	-2.26	-2.63	-3.61	-3.03	-15.0
E	-3.64	-15.0	-3.51	-2.65	-3.39	-3.41	-3.21	-2.63	-15.0
F	-1.89	-15.0	-2.35	-2.5	-1.34	-2.43	-2.18	-1.71	-15.0
G	-2.32	-15.0	-3.04	-2.63	-2.56	-2.3	-3.13	-2.96	-15.0
H	-2.67	-15.0	-2.58	-2.58	-2.05	-3.32	-3.13	-2.16	-15.0
I	-1.65	-2.55	-2.8	-3.44	-2.74	-2.79	-2.2	-2.69	-2.1
K	-2.51	-15.0	-3.65	-2.93	-3.34	-3.77	-2.97	-3.27	-15.0
L	-2.32	-1.7	-2.09	-2.49	-2.71	-2.63	-2.62	-2.01	-2.74
M	-0.39	-1.39	-1.79	-3.01	-3.43	-1.38	-1.33	-0.97	-2.96
N	-3.12	-15.0	-3.31	-2.22	-2.36	-2.3	-3.14	-3.31	-15.0
P	-3.61	-15.0	-2.97	-2.64	-2.42	-2.31	-1.83	-2.42	-15.0
Q	-2.76	-15.0	-2.81	-2.63	-3.06	-2.84	-2.12	-3.05	-15.0
R	-1.92	-15.0	-3.41	-2.61	-3.05	-3.76	-3.43	-3.02	-15.0
S	-2.39	-15.0	-2.04	-2.12	-2.83	-3.04	-2.73	-2.02	-15.0
T	-2.89	-3.58	-2.6	-2.48	-2.17	-2.58	-2.67	-3.14	-3.7
V	-2.44	-2.64	-2.68	-3.29	-2.49	-2.25	-2.68	-2.8	-1.7
W	-0.14	-15.0	-1.01	-2.94	-1.77	-2.77	-2.85	-2.13	-15.0
Y	-1.46	-15.0	-1.67	-2.7	-1.92	-2.39	-1.35	-3.37	-15.0

Coefficients of the polynomial method. For each of the 20 amino acid residues, the amount they contribute to the polynomial score is shown (for all nine positions in which they may occur).

from a training set of 161 peptides by a method analogous to that used by epidemiologists to calculate risk factors for developing a disease. All peptides in the training set contain the canonical motif for HLA-A2.1, so that they all contain the “correct” residue at the anchor positions (2 and 9). For i other than 2 and 9 (for the non-anchor positions), the average negative \log_{10} of IC_{50} of all the peptides carrying R at position i is calculated and used as the estimate of R_i . This is a slightly modified version of the method reported by Ruppert *et al.*^[9]. The values of the R_i terms of HLA-A2 specificity are shown in Table 1.

To calculate the polynomial method score of peptides in this study, the R_i values corresponding to the sequence of the given peptide were added together. If this sum exceeded a chosen threshold, the peptide was predicted to bind. The threshold was chosen as the number that gave a relatively clean separation between binders and non-binders in the training set. In the present study the threshold was -24.

Prediction result analysis

The previously reported CTL epitope of the antigens investigated in this study were obtained from the “Find Your Motif” section of SYFPEITHI database. These epitopes were screened out from the primary predicted epitopes.

RESULTS

SYFPEITHI prediction

The SYFPEITHI prediction values of all possible nonamers of a given protein sequence were added together and the ten high-scoring peptides of each protein were selected for further analysis. Scores for all predicted epitopes specifically for HLA-A*0201 are given in Table 2. The predicted epitopes were accorded with the simple motif (SM) of HLA-A2.1 (the presence of L, M or I at position 2 and L, V or I at position 9). The SYFPEITHI predicted scores of these epitopes were usually higher than 20.

Polynomial method analysis

The primary predicted epitopes were compared with epitopes that were demonstrated in previous research. Six reported HLA-A2 restricted CTL epitopes (MAGE-1₂₇₈₋₂₈₆ KVLEYVIVK^[10], MAGE-3₁₁₂₋₁₂₀ KVAELVHFL^[11], MAGE-3₂₇₁₋₂₇₉ FLWGPRLV^[12], P53₁₈₇₋₁₉₇ GLAPPQHLIRV^[13], P53₃₂₂₋₃₃₀ PLDGEYFTL^[14], AFP₁₅₈₋₁₆₆ FMNKFYIEI^[15]) were eliminated from the SYFPEITHI prediction results. The polynomial scores of predicted epitopes are shown in Table 2. The epitopes which polynomial scores less than -24 (MAGE-1_{101-109, 301-309, 15-23}, MAGE-8₂₀₋₂₈, P53_{129-137, 193-201, 256-264, 113-121}, AFP₄₁₀₋₄₁₈) were eliminated from the prediction results. At last, thirty-five epitopes were selected from the predicted results for further study.

Table 2 HLA-A2 restricted epitope prediction results

Antigens	AApos	Sequence	SYFPEITHI score	Polynomial score
MAGE-1 (MAG1)	194	FLIIVLVMI	27	-17.07
	38	LVLGTLEEV	26	-20.77
	278	KVLEYVIVK	26	-21.23
	101	VITKKVADL	25	-24.23
	301	ALREEEEGV	25	-24.81
	187	QIMPKTGFL	24	-23.24
	93	ILES LFRV	23	-22.32
	105	KVADLVGFL	23	-23.12
	15	ALEAQVEAL	22	-25.17
	89	STSCILES L	22	-23.31
MAGE-3 (MAG3)	108	ALSRKVAEL	31	-22.04
	201	LLIIVLAI	28	-22.52
	200	GLLIIVLAI	27	-22.33
	271	FLWGPRLV	27	-19.47
	220	KIWEELS VL	26	-23.01
	112	KVAELVHFL	25	-23.14
	237	SILGDPKKL	25	-22.73

	176	YIFATCLGL	23	-21.73
	238	ILGDPKKLL	23	-23.01
	174	HL YIFATCL	22	-21.30
MAGE-8 (MAG8)	111	ALDEKVAEL	33	-23.5
	45	LIMGTLEEV	29	-21.63
	204	LLIIVLGM I	26	-21.58
	115	KVAELVRFL	24	-23.44
	179	YILVTCLGL	24	-22.10
	71	SLTVDSTL	23	-23.39
	203	GLLIIVLGM	23	-23.08
	205	LIIVLGMIL	23	-22.73
	<u>20</u>	<u>GEAPGLMDV</u>	<u>21</u>	<u>-34.01</u>
	25	LMDVQIPTA	21	-23.75
P53 (P53)	24	KLLPENNVL	26	-23.31
	187	GLAPPQH LI	25	-21.96
	264	LLGRNSFEV	24	-21.58
	<u>129</u>	<u>ALNKMFCQL</u>	<u>23</u>	<u>-24.94</u>
	132	KMFCQLAKT	22	-23.22
	<u>193</u>	<u>HLIRVEGNL</u>	<u>22</u>	<u>-24.86</u>
	65	RMPEAAPPV	21	-20.47
	322	PLDGEYFTL	21	-25.24
	<u>256</u>	<u>TLEDSSGNL</u>	<u>20</u>	<u>-25.86</u>
	<u>113</u>	<u>FLHSGTAKS</u>	<u>19</u>	<u>-34.05</u>
AFP (FETA)	90	QLPAFLEEL	27	-22.64
	257	KLVLDDVAHV	27	-20.10
	158	FMNKFIYEI	26	-20.37
	172	FLYAPTILL	25	-19.87
	410	ALAKRSCGL	25	-24.57
	218	LLNQHACAV	24	-22.45
	242	KLSQKFTKV	24	-22.29
	571	VIADFSGLL	23	-22.76
	599	LISKTRAAAL	23	-23.93
	8	FLIFLLNFT	22	-22.78

A: the epitopes were reported; A: the prediction epitopes whose polynomial scores were less than -24.

DISCUSSION

The induction of antigen-specific cytotoxic T lymphocytes (CTL) has been suggested to be efficacious in the prevention and treatment of various types of tumor. Determination of peptides that elicit a T-cell response *in vivo* is important for identifying autoimmune and CTL epitopes and for peptide vaccine design. About 120 CTL epitopes presented by HLA-A, B, C molecules have been reported in recent study^[4] and some epitopes have been used as peptide vaccine in animal and clinical experiments^[16,17].

While sufficient conditions for a peptide to be a T-cell epitope are not well known, one well established necessary condition is that they bind to MHC molecules. Hence considerable effort has been made to measure the binding affinities of peptides to MHC molecules. A reliable theoretical method that rapidly predict whether a peptide bind is of great practical utility. The methods of epitope prediction were constructed by immunologists in recent decades. These predictions are based on comparisons of precursor peptide sequences known to contain T cell epitopes. It is the discovery of allele-specific motifs shared by eluted natural MHC ligands that finally allow the exact prediction of peptides from a given protein sequence presented on MHC class I molecules. Currently, several algorithms are publicly available for predicting the HLA-binding affinities of peptides, such as the SYFPEITHI database (<http://uni-tuebingen.de/uni/kxi/>)^[18]. SYFPEITHI uses motif matrices

deduced from refined motifs based on the pool sequence and single peptide analysis exclusively of natural ligands. Potential binders for various MHC class I molecules are ranked according in the presence of primary and secondary anchor amino acids as well as favored and disfavored amino acids.

However, motifs are the simplest independent binding site models: they specify positions with required residues and certain other positions with prohibited residues. One way to include more detail than motifs in an independent binding model is the polynomial method. Polynomial method is based on statistical parameter estimation assuming independent binding of the side-chains of residues^[9]. Its sensitivity and positive predictive value are better than the simple motif (such as syfpeithi database)^[19]. This and other^[20-22] independent binding site methods that assign a score to a peptide have been called quantitative motifs^[23]. Threshold is an integral part of the polynomial method. In the above analysis the threshold was set to -24. In this study, 35 candidate epitopes were selected for further immunology experiments by the SYFPEITHI prediction combined with polynomial methods.

Hepatocellular carcinoma (HCC) is one of the most common neoplasms worldwide. It is the major cause of death in China and some regions of Asia. The pathogenic mechanisms responsible for HCC are not well defined, and therapeutic means, especially in unresectable HCCs, are still unsatisfactory. In Chinese HCC patients, the proportion of HLA-A2 is higher than other alleles^[8]. Therefore, identification of HLA-A2 restricted CTL epitopes of tumor antigen closely associated with HCC is very valuable for HCC immunotherapy in China. In this study, the antigens of MAGE gene family (MAGE-1, MAGE-3 and MAGE-8), P53 mutation gene and AFP were selected as objective proteins because these antigens have been demonstrated to be closely associated with HCC^[24-26,15]. The candidate HLA-A2 restricted CTL epitopes of these antigens were predicted using the syfpeithi prediction combined with polynomial methods.

The epitope prediction has been carried in tumor antigen specificity for CTL epitope identification in most studies recently. Schirle *et al*^[14] reported that two new CTL epitopes of gastrointestinal tumor were identified by epitope prediction combined with acid eluted methods. Epitope prediction also has been combined with epitope reconstruction and immunology assay by Pascolo *et al*^[10] and a new CTL epitope of MAGE-1 was identified.

If immunological activity of the candidate epitopes has been identified by further immunology assay, most of the patients with HCC could be potentially candidates for specific immunotherapy against these epitopes, including vaccination using APC pulsed with HLA-A2-restricted CTL epitopes or the adoptive transfer of specific CTL generated from PBMC by stimulation with the epitopes. The results of this study will benefit to most patients with HCC in the future.

REFERENCES

- 1 Dong HL, Sui YF, Li ZS, Qu P, Wu W, Ye J, Zhang XM, Lu SY. Efficient induction of cytotoxic T lymphocytes specific to hepatocellular carcinoma using HLA-A2-restricted MAGE-n peptide *in vitro*. *Cancer Lett* 2004; **211**: 219-225
- 2 van der Bruggen P, Traversari C, Chomez P, Lurquin C, De Plaen E, Van den Eynde B, Knuth A, Boon T. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science* 1991; **254**: 1643-1647
- 3 Katsura F, Eura M, Chikamatsu K, Oiso M, Yumoto E, Ishikawa T. Analysis of individual specific cytotoxic T lymphocytes for two MAGE-3-derived epitopes presented by HLA-A24. *Jpn J Clin Oncol* 2000; **30**: 117-121
- 4 Renkvist N, Castelli C, Robbins PF, Parmiani G. A listing of

- human tumor antigens recognized by T cells. *Cancer Immunol Immunother* 2001; **50**: 3-15
- 5 **Gebo KA**, Chander G, Jenckes MW, Ghanem KG, Herlong HF, Torbenson MS, El-Kamary SS, Bass EB. Screening tests for hepatocellular carcinoma in patients with chronic hepatitis C: a systematic review. *Hepatology* 2002; **36**: S84-S92
 - 6 **Butterfield LH**, Ribas A. Immunotherapy of hepatocellular carcinoma. *Expert Opin Biol Ther* 2002; **2**: 123-133
 - 7 **Ladhams A**, Schmidt C, Sing G, Butterworth L, Fielding G, Tesar P, Strong R, Leggett B, Powell L, Maddern G, Ellem K, Cooksley G. Treatment of non-resectable hepatocellular carcinoma with autologous tumor-pulsed dendritic cells. *J Gastroenterol Hepatol* 2002; **17**: 889-896
 - 8 **Ming LH**, Zhu YR, Harris CC, Wang XH, Wu ZY, Sun ZT. Association of HLA-A02 genotype with hepatocellular carcinoma in a prevalent area of China. *Chin J Immunol* 1999; **15**: 305-310
 - 9 **Ruppert J**, Sidney J, Celis E, Kubo RT, Grey HM, Sette A. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* 1993; **74**: 929-937
 - 10 **Pascolo S**, Schirle M, Guckel B, Dumrese T, Stumm S, Kayser S, Moris A, Wallwiener D, Rammensee HG, Stevanovic S. A MAGE-A1 HLA-A*0201 epitope identified by mass spectrometry. *Cancer Res* 2001; **61**: 4072-4077
 - 11 **Kawashima I**, Hudson SJ, Tsai V, Southwood S, Takesako K, Appella E, Sette A, Celis E. The multi-epitope approach for immunotherapy for cancer: identification of several CTL epitopes from various tumor-associated antigens expressed on solid epithelial tumors. *Hum Immunol* 1998; **59**: 1-14
 - 12 **van der Bruggen P**, Bastin J, Gajewski T, Coulie PG, Boel P, De Smet C, Traversari C, Townsend A, Boon T. A peptide encoded by human gene MAGE-3 and presented by HLA-A2 induces cytolytic T lymphocytes that recognize tumor cells expressing MAGE-3. *Eur J Immunol* 1994; **24**: 3038-3043
 - 13 **Theobald M**, Biggs J, Hernandez J, Lustgarten J, Labadie C, Sherman LA. Tolerance to p53 by A2.1-restricted cytotoxic T lymphocytes. *J Exp Med* 1997; **185**: 833-841
 - 14 **Schirle M**, Keilholz W, Weber B, Gouttefangeas C, Dumrese T, Becker HD, Stevanovic S, Rammensee HG. Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *Eur J Immunol* 2000; **30**: 2216-2225
 - 15 **Butterfield LH**, Meng WS, Koh A, Vollmer CM, Ribas A, Dissette VB, Faull K, Glaspy JA, McBride WH, Economou JS. T cell responses to HLA-A*0201-restricted peptides derived from human alpha fetoprotein. *J Immunol* 2001; **166**: 5300-5308
 - 16 **Thurner B**, Haendle I, Röder C, Dieckmann D, Keikavoussi P, Jonuleit H, Bender A, Maczek C, Schreiner D, von den Driesch P, Bröcker EB, Steinman RM, Enk A, Kämpgen E, Schuler G. Vaccination with mage-3A1 peptide-pulsed mature, monocyte-derived dendritic cells expands specific cytotoxic T cells and induces regression of some metastases in advanced stage IV melanoma. *J Exp Med* 1999; **190**: 1669-1678
 - 17 **Dagher R**, Long LM, Read EJ, Leitman SF, Carter CS, Tsokos M, Goletz TJ, Avila N, Berzofsky JA, Helman LJ, Mackall CL. Pilot trial of tumor-specific peptide vaccination and continuous infusion interleukin-2 in patients with recurrent Ewing sarcoma and alveolar rhabdomyosarcoma: an inter-institute NIH study. *Med Pediatr Oncol* 2002; **38**: 158-164
 - 18 **Rammensee H**, Bachmann J, Emmerich NP, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 1999; **50**: 213-219
 - 19 **Gulukota K**, Sidney J, Sette A, DeLisi C. Two complementary methods for predicting peptides binding major histocompatibility complex molecules. *J Mol Biol* 1997; **267**: 1258-1267
 - 20 **Marshall KW**, Wilson KJ, Liang J, Woods A, Zaller D, Rothbard JB. Prediction of peptide affinity to HLA DRB1*0401. *J Immunol* 1995; **154**: 5927-5933
 - 21 **Parker KC**, Bednarek MA, Coligan JE. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol* 1994; **152**: 163-175
 - 22 **Hammer J**, Bono E, Gallazzi F, Belunis C, Nagy Z, Sinigaglia F. Precise prediction of major histocompatibility complex class II-peptide interaction based on peptide side chain scanning. *J Exp Med* 1994; **180**: 2353-2358
 - 23 **Hammer J**. New methods to predict MHC-binding sequences within protein antigens. *Curr Opin Immunol* 1995; **7**: 263-269
 - 24 **Tahara K**, Mori M, Sadanaga N, Sakamoto Y, Kitano S, Makuuchi M. Expression of the MAGE gene family in human hepatocellular carcinoma. *Cancer* 1999; **85**: 1234-1240
 - 25 **Kobayashi Y**, Higashi T, Nouse K, Nakatsukasa H, Ishizaki M, Kaneyoshi T, Toshikuni N, Kariyama K, Nakayama E, Tsuji T. Expression of MAGE, GAGE and BAGE genes in human liver diseases: utility as molecular markers for hepatocellular carcinoma. *J Hepatol* 2000; **32**: 612-617
 - 26 **Charuruks N**, Tangkijvanich P, Voravud N, Chatsantikul R, Theamboonlers A, Poovorawan Y. Clinical significance of p53 antigen and anti-p53 antibodies in the sera of hepatocellular carcinoma patients. *J Gastroenterol* 2001; **36**: 830-836

Edited by Wang XL and Xu FM