*ISSN (e): **2250 – 3005** || Volume, 07 || Issue, 06|| June – 2017 ||*
International Journal of Computational Engineering Research (IJCER)

Generation of tissue-specificand promiscuous HL Aligand databasesusing DNA microarrays and virtual HLA class II matrices

Meenakshi Panda, Rupasmita Sahoo
Gandhi Institute of Excellent Technocrats, Bhubaneswar,India
Rayagada Institute of Technology and Management, Rayagada, Odisha, India

ABSTRACT
Most pockets in the human leukocyte antigen–group DR (HLA-DR) groove are shaped by clusters ofpolymorphic residues and, thus, have distinct chemical and size characteristics in different HLA-DR alle-les. Each HLA-DR pocket can be characterized by "pocket profiles," a quantitative representation of theinteraction of all natural amino acid residues with a given pocket. In this report we demonstrate thatpocket profiles are nearly independent of the remaining HLA-DR cleft. A small database of profiles wassufficient to generate a large number of HLA-DR matrices, representing the majority of human HLA-DRpeptide-binding specificity. These virtual matrices were incorporated in software (TEPITOPE) capable ofpredictingpromiscuousHLAclassIIligands.Thissoftware,incombinationwithDNAmicroarraytechnol-
ogy,hasprovidedanewtoolforthegenerationofcomprehensivedatabasesofcandidatepromiscuousT-cell epitopes in human disease tissues. First, DNA microarrays are used to reveal genes that are specifi-cally expressed or upregulated in disease tissues. Second, the prediction software enables the scanningof these genes for promiscuous HLA-DR binding sites. In an example, we demonstrate that starting fromnearly 20,000 genes, a database of candidate colon cancer–specific and promiscuous T-cell epitopescould be fully populated within a matter of days. Our approach has implications for the development ofepitope-basedvaccines.

Keywords:HLAmatrix,epitopeprediction,DNAmicroarrays,tumorimmunology,genomics

Helper T-cell activation is essential for the initiation of a protectiveimmune response to pathogens and tumors [1,2]. Human leukocyteantigen–groupDR(HLA-DR),thepredominantisotypeofthehuman class II major histocompatibility complex (MHC), plays acentral role in helper T-cell selection and activation. Proteins ofHLA-DR bind peptide fragments derived from protein antigens anddisplay them on the surface of antigen-presenting cells for interac-tionwithantigen-specificreceptorsofTlymphocytes[1].
X-ray crystallographic studies demonstrated that the HLA-DRligand bindinggrooveconsistsofpockets,resultinginstrongprefer-ences for interaction with particular amino acid side chains of theligands[3–6].MoleculesofHLA-DRareextremelypolymorphic.Polymorphic residues are often involved in forming HLA-DR pock-ets;consequently,pocketsofdifferentHLA-DRallelescan beofdis-tinctchemicalandsizecharacteristics.Someoftheligandsidechainsinteract with the pockets and increase the overall binding affinityandspecificity of ligands, whereas others interfere withpocketresidues and reduce binding[7]. Therefore, the pocket specificity canbe characterized either topographically (i.e., by differences in theamino acid residues forming the pockets) or functionally (i.e., bysubstituting the corresponding peptide ligand position with all nat-ural amino acid residues and by quantifying their effects on binding["pocket profiles"]). The sum of all pocket profiles of a given HLA-DRalleleisdefinedasa"quantitativematrix"[8].

We and others have demonstrated that matrices are powerfultoolstopredictHLAclassIIligands[8,9].Incontrasttopreviousall-or-nothing rules and approaches that are based on artificial neural net-works[10,11],matrix-basedpredictionsrelyonmathematicalprocessingof individual peptide side chain effects (see ref. 12 for a detailedcomparison of bioinformatic tools used for HLA class II ligand pre-diction). A typical matrix-based algorithm first extracts all possiblepeptide frames from a given protein sequence. Subsequently, thecorresponding position- and amino acid–specific matrix values areassigned to each residue of these peptide frames. Finally, the sum ofthese matrix values is determined for each frame. It has been shownthat the resulting numerical values ("peptide scores") correlate withthebindingaffinityofHLA-DRligands,thusmakingmatricesimportanttoolsforthepredictionofcandidateT-cellepitopes[13,14].

HLA-DRmoleculesaccountformorethan90%oftheHLAclassIIisotypesexpressedonantigen-presentingcells.AlthoughtheHLA-DRAlocusismonomorphic,morethan100 alleleshavebeendescribedfortheHLA-DRB1locus[15].Matriceshavesofarbeendeterminedbymeasuring allpossiblepocketprofiles ona givenHLA-DRallele.Hence,thedeterminationofasingleHLA-DRmatrixrequiredhundreds of individualpeptides andthousands ofpeptide binding assays[13]; a global coverage of HLA class II bindingspecificity seemed,therefore,unlikely inthenearfuture. Inthisreport,wedemonstratethatpocketprofilesarenearlyindependent of the remaining HLA-DR groove. Thus, once a pocket profile hasbeen determined in vitro, it can be shared among other HLA-DRalleles as long as their amino acid residues contributing to the pock-et are identical. Consequently, a relatively small number of pocketprofiles can be assigned to a large number of HLA-DR alleles viasequencecomparison.Theresultingvirtualmatricescoverthemajorityofhuman HLA-DRspecificity.

A comprehensive database of candidate promiscuous T-cell epi-topes in tumors or pathogens would be of great value for vaccinestrategies. Major bottlenecks so far have included not only the needto determine quantitative matrices for each polymorphic HLA-DRallele, but also the lack of gene expression data enabling, for exam-ple, a comprehensive selection of genes expressed in disease but notin normal tissue. The latter has become feasible by the recent devel-opment of DNA microarray technology[16,17]: DNA microarrays areused to monitor and compare the expression of thousands of genessimultaneouslyandarethuscapableofidentifyinglargepoolsofdif-ferentially expressed candidate antigens. The former is resolved inthisstudybyapplyingtheaboveconceptofvirtualHLAmatrices.

RESULTS AND DISCUSSION

Allele independence of pocket profiles. The value of matrix-basedcomputational algorithms for the prediction of helper T-cell epi-topes has been demonstrated beyond doubt, as exemplified by therecent discovery of a human leukocyte function-associated antigen-1(LFA-1)peptideasthecandidateautoantigeninLymearthritis[18]orbyarecentx-raycrystalstructureofaDRB1*0401–collagenIIpep-tidecomplex[6,19].BoththeLFA-1andthecollagenpeptideswereiden-tified using our previously described DRB1*0401 matrix[8]. In thisreportweproposeanewstrategy(Fig.1A)thatleadstobothabroadcoverage of human HLA-DR binding specificity and the possibilityofcreatinggenomicdatabasesofcandidateT-cellepitopes.

Wehavepreviouslydemonstratedthatpocketspecificityprofilesaremostlyindependentofneighboringligandsidechains[8]. Moreover,initialanalyseson DRB1*04subtypeshavealsosuggestedthatpocketprofilesmightbeindependentoftheremainingHLA-DRgroove[19].Obviously,thelatterwouldhaveimportantimplicationsinthat pocket profiles are only determined once and can subsequentlybesharedamongallelesaslongastheyarepredictedtohavesimilar

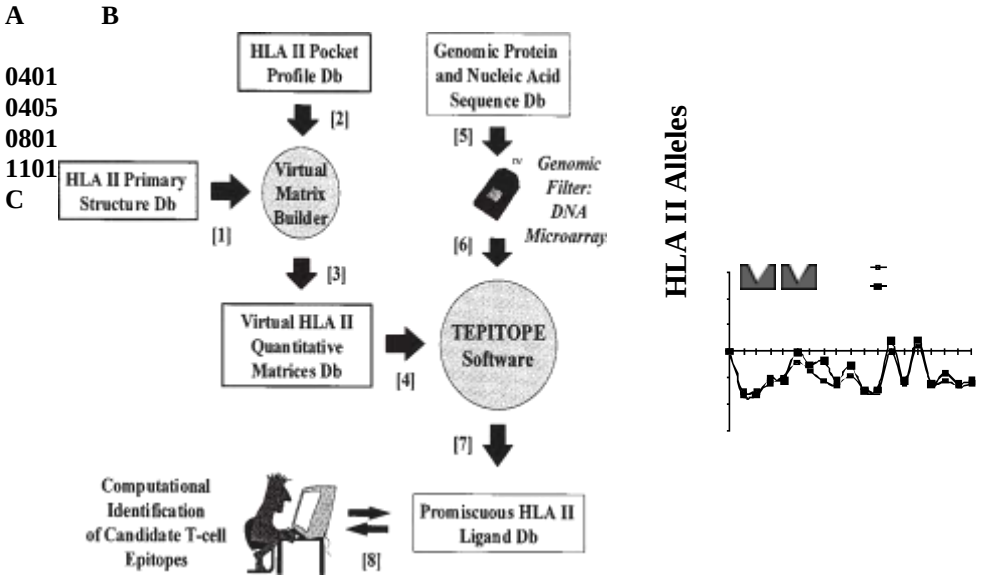
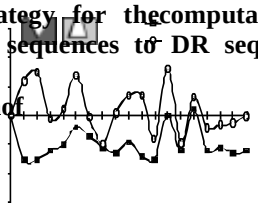


Figure1.Alleleindependenceofpocketprofilesleadstowidecoverage

1A DEFGHIKLMN PQRSTVWY
0.1
0.01
0.001
1
ADEFGHIKLMNPQRSTVWY
0.1
0.01
0.001

of HLA-DR binding specificity. (A) Overview of the strategy for thecomputationalgenerationofHLA-DRliganddatabases:(1)Useofmultiple alignment of HLA-DR sequences to DR sequences of knownthree-dimensionalstructuretoassignpolymorphicresiduestoHLA-DRpockets;classificationofpocketsaccordingtotheircompositionof



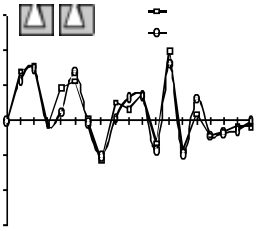
1000
100
10
1
Pocket9

DRB1*0401DRB1*0801

1000
100
10
1
Pocket9

DRB1*0405DRB1*0801

polymorphicresidues.(2)Determinationofadatabaseofpocketspecificityprofiles,asdescribedintheExperimentalprotocol.(3)
ADEFGHIKLMNPQRSTVWY
0.1
ADEFGHIKLMNPQRSTVWY
0.1



Combination of the data coming from (1) and (2) for the assignment ofprofiles to corresponding pockets, based on the demonstration thatpocketssharingthesamepolymorphicresiduesexhibitasimilarpocketspecificityprofile;assemblyofvirtualHLA-DRmatricesusingtheassignedpocketprofiles.(4)IncorporationoftheobtainedHLA-DR
0.01
0.001

0.01
0.001

Peptideaminoacids

virtual matrix database into an epitope prediction software. (5) Presentation of protein/gene/EST sequence databases on DNA microarrays. (6)Identification of specifically expressed or upregulated genes in disease tissues by DNA microarray expression mapping. (7) Scanning of theidentifiedsequencesusingthepredictionsoftware,allowingtheidentificationofcandidatepromiscuousHLA-DRligands.(8)Useofthegenerated HLAclassIIliganddatabasefortheidentificationofcandidatepromiscuoushelperT-cellepitopes.(B)Schematicrepresentationofthemodularstructure of the HLA-DR binding groove. The binding clefts of four HLA-DR allotypes are compared. The cleft regions 1–3 are constituted bymonomorphicresiduesmostlycomingfromtheDRαchain(exceptforonedimorphicresiduefromtheDRβchain(Gly/Val86),composingpocket1); positions 5 and 8 were excluded because peptide side chains at these positions are oriented away from the DR binding cleft, as shown incrystal structure analyses[3,5,6]; pockets 4, 6, 7, and 9 are mainly formed by DR β chain polymorphic residues and are responsible for the allelespecificity of HLA-DR–ligand interaction. The modular structure of the HLA-DR binding groove enables the free exchange of functional pocketprofiles, as long as the polymorphic residues forming the pockets are the same. (C) Pockets on different alleles sharing the same polymorphicresiduesexhibitsimilarpocketspecificityprofiles.Comparisonofpocketspecificityprofilesobtainedforpocket 9fromHLA-DRalleles,whichareformed either by identical polymorphic residues (top left and bottom right panels) or by different ones (top right and bottom left panels),demonstratingthatHLA-DRprimarystructurehomologycanbesufficienttoassigndefinedbindingspecificityprofilestogivenpockets.Pocketspecificity profiles were determined in HLA-DR competitive binding assays by quantifying the effects of all amino acid substitutions at a givenpositionofaffinityoptimized,alanine-baseddesignerpeptides,asdescribedintheExperimentalProtocol.RelativebindingvalueswerecalculatedbynormalizingexperimentalIC_{50}datawiththeIC_{50}valueobtainedforalanineatthesamepeptideposition(IC_{50}Ala/IC_{50}substitution).

table1.Pocketprofiledatabase.

DRsource	Pocket Profile ID#		Polymorphicpocket residues(β-chain)	A	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
B1*0101	4	1	13F;70Q;71R;74A;78Y	0	-2.4	-0.4	0.08	-0.7	-0.7	0.5	-2.1	0.9	0.8	0.04	-1.9	0.1	-2.1	-0.7	-1	-0.05	-1.8	-1.1
B1*1501	4	2	13R;70Q;71A;74A;78Y	0	-0.4	-0.6	2.4	0	1.1	0.6	-0.7	0.5	1	-0.2	-0.3	-0.8	0.2	-0.3	-0.3	0.2	0.4	2.5
B1*0301	4	3	13S;70Q;71K;74R;78Y	0	2.3	-1	-1	0.5	0	0.5	-1	0	0	0.2	-1	0	-1	0.7	-1	0	-1.0	-1
B1*0401	4	4	13H;70Q;71K;74A;78Y	0	1.4	1.5	-0.9	-1.6	1.1	0.8	-1.7	0.8	0.9	0.9	-1.6	0.8	-1.9	0.8	0.7	-0.9	-1.2	-1.6
B1*0402	4	5	13H;70D;71E;74A;78Y	0	-2.3	-2.3	0.3	-0.7	1.2	0.08	0.1	-0.6	0.6	-0.4	-1.3	-0.4	1	-1	-0.5	-0.7	1.6	-0.4
B1*0404	4	6	13H;70Q;71R;74A;78Y	0	-1.1	-1.1	1	-2.4	-1	1.1	-1.5	1	1.8	-0.7	-1.3	0	-2.4	-0.7	-0.9	0.5	-0.05	-0.4
B1*1101	4	7	13S;70D;71R;74A;78Y	0	-1.7	-1.7	0.4	-1.7	-0.6	0.9	-0.5	1.1	1	0	-1.7	-0.4	-0.7	-0.7	-0.6	0.4	-0.1	-0.7
B1*0701	4	8	13Y;70D;71R;74Q;78V	0	-1.6	-1.4	0.2	-1.1	0.1	1.1	-1.3	-0.8	-0.4	-1.1	-1.2	-1.5	-1.1	1.5	1.4	0.9	-1.1	-0.9
B1*0801	4	9	13G;70D;71R;74L;78Y	0	-1	-1	0.5	-1	-1	0.3	2.3	0.7	1.4	0	-1	-1	2.3	-1	-1	0.3	0	2.2
B5*0101	4	10	13Y;70D;71R;74A;78Y	0	-1.9	-1.3	-0.6	-1.6	-1.4	1.3	-1.7	0.6	1.7	-1.7	-1.5	-0.7	-1.7	-0.5	0.3	1.1	-1.4	-0.6
B1*1302	4	11	13S;70D;71E;74A;78Y	0	-1.4	-1.1	0.8	-1.5	1.5	-0.6	0.8	0.4	0.8	0.1	-1.5	0.6	0.2	-0.6	-1.1	-0.9	0.7	0.4
B1*0101	6	1	11L	0	-2.7	-2.4	-2.1	-0.3	-2.2	-1.9	-2	-2	-1.8	-1.1	-0.2	-1.8	-1.8	-0.6	-1.2	-1.1	-2.4	-2
B1*1501	6	2	11P	0	-0.4	-1	-0.3	0.5	-0.5	0.05	-0.3	0.2	0.1	0.7	-0.2	-0.8	1	0.6	-0.04	-0.3	-0.4	0.4
B1*0801	6	3	11S	0	-2.4	-1.4	-1.4	-0.7	-0.1	0.7	1.3	0.2	-0.9	-0.6	0.5	-0.3	1	-0.1	0.8	1.2	-1.4	-1.4
B1*0404	6	4	11V	0	-1.1	-2.4	-1.1	-1.5	-1.4	-0.1	-2.4	-1.1	-1.1	1.3	0	-1.5	-2.4	1	1.9	0.9	-1	-1.5
B1*0701	6	5	11G	0	-2.5	-2.5	-0.8	-0.6	-0.8	-0.5	-1.1	-0.9	-0.8	-0.6	-0.5	-1.1	-1.1	0.6	-0.08	0.1	-0.9	-1

B5*0101	6	6	11D	0	-2	-2	-1.7	-0.3	-1.2	-1.4	-1.5	-1	-1.5	-1.3	0.2	-1.4	-1.3	-0.5	-0.8	-1.3	-1.7	-1
B1*0101	7	1	28E;30C;47Y;60 1W;67L;71R		-2	-0.6	0.3	-1.1	0.1	0.6	-0.2	0.3	0.09	0.1	0.07	0.2	0.09	-0.2	0.09	0.7	-0.08	0.5
B1*1501	7	2	28D;30Y;47F;60 1W;67I;71A		-0.7	-0.7	1.4	0	0.6	1.5	-0.3	1.9	1.7	0.7	0.3	-0.3	-0.5	0.3	0.2	0.3	0.6	0.7
B1*0301	7	3	28D;30Y;47F;60 1W;67L;71K		-0.6	-0.2	0.5	0.1	-0.8	0.4	-0.9	0.2	1.1	-0.09	0.7	-0.1	-0.9	0.07	-0.1	0.2	-0.6	-0.05
B1*0401	7	4	28D;30Y;47Y;60 1W;67L;71K		-0.3	0.2	-1	-1.3	0	0.08	-0.3	0.7	0.8	0.6	-0.7	0	-1.2	-0.2	-0.1	0.08	-1.4	-1.2
B1*0402	7	5	28D;30Y;47Y;60 1W;67I;71E		-2.1	-1.2	0.5	-2.1	0.5	0.5	0	1	0.8	0.6	-1	1.1	1.7	-0.4	0.1	0.2	1.4	0.9
B1*0404	7	6	28D;30Y;47Y;60 1W;67L;71R		-1.2	-0.7	-0.05	-1.2	-0.4	0.08	-1.3	0.3	0.7	0.7	-1	-0.2	-0.9	0.5	0.4	-0.1	-0.7	-0.2
B1*1101	7	7	28D;30Y;47F;60 1W;67F;71R		-2.7	-1.3	-0.4	-0.4	-0.2	0.8	-0.2	1.5	1.3	-1	-0.05	-1.1	-0.4	-1.3	-1.3	-0.6	-0.4	-0.4
B1*0701	7	8	28E;30L;47Y;60 1W;67L;71R		-1.3	0.9	2.1	0	0.9	2.4	0.5	2.2	1.8	1.4	-0.2	1.1	0.7	0.4	0.9	1.6	1.4	1.7
B1*0801	7	9	28D;30Y;47Y;60 1W;67F;71R		-2.4	-2.4	-0.9	-0.5	-0.9	-0.8	-0.8	-0.3	-0.3	-1.3	-1.2	-1.2	-0.6	-1.3	-2.4	-1.1	-1.2	-1.1
B5*0101	7	10	28H;30D;47Y;60 1W;67F;71R		-1.5	-0.9	1.5	0.6	1.2	1.2	0.9	0.6	0.4	0.5	-0.6	0.7	1.3	-0.2	0.3	-0.3	0.4	1.2
B1*1302	7	11	28D;30Y;47F;60 1W;67I;71E		-1.5	-1	0.6	-1.5	0.3	-0.5	0	0.4	0	0.1	-0.5	-0.4	1.2	-0.9	-0.9	-0.1	0.4	-0.2
B1*0101	9	1	9W;37S;57D;60 Y;61W		-1.9	-1.9	-0.4	-0.8	-1.1	0.7	-1.7	0.5	0.08	-1.2	-1.1	-1.6	-1	-0.3	-0.2	0.3	-1.4	-0.9
B1*0301	9	2	9E;37N;57D;60 Y;61W		-0.6	-0.3	0.9	0.4	-0.5	0.6	-0.2	-0.04	1.1	-0.6	-0.3	-0.2	0.5	1.1	-0.5	0.3	-1	0.3
B1*1101	9	3	9E;37Y;57D;60 Y;61W		-1.7	-1.7	-1	-1	0.08	-0.3	-0.3	-1	-0.4	-1.4	-1.3	0.5	-1	0.7	-1.2	-0.7	-1	-1
B1*0701	9	4	9W;37F;57V;60 S;61W		-1.2	-0.3	2.1	-0.6	-0.2	3.4	-1.1	3.4	2	-0.5	-0.6	-0.9	-0.8	-0.3	0.4	2	0.8	1.1
B1*0801	9	5	9E;37Y;57S;60 Y;61W		1	1.3	-0.1	0.3	1.3	-0.1	-1	0	0.7	0.6	-0.9	1.3	-1	0.7	-0.3	-0.4	-0.1	0.1
B5*0101	9	6	9Q;37D;57D;60 Y;61W		-1.5	-0.6	1.2	0.4	1	1.2	2.7	1.3	0.5	0	-0.8	0.7	2.5	0.7	-0.2	-0.2	-0.7	1.3

Aminoacidresidue

ThistableshowstheHLA-DRallelesusedtogeneratethedifferentprofilesforthepolymorphicpockets4,6,7,and9.Differentprofilesforeachpocketareindicatedwithidentificationnumbers.Foreachspecificprofilethe DR β-chain polymorphic residues composing each pocket have been specified and the relative values of the effects of all natural amino acid side residues at each position are reported. Data are expressed as the log-arithmofthealanine-normalizedrelativebindingdatacalculatedasinFigure1C.

pockettopographies(Fig.1B).Totestthishypothesis,wedetermined pocketspecificityprofilesforseveralHLApock-ets and compared them with each other (Fig. 1C). The align-mentofHLA-DRsequenceswithDRsequencesofknownthree-dimensionalstructures[3–6]indicatedthatthepolymor-phicresiduesconstitutingpocket9inDRB1*0401andDRB1*1101allotypesareidentical.Thisfindingisconsistentwiththeobservationthatthe pocketprofilesforbothallelesaresimilar(Fig.1C).Incontrast,thealignmenttoHLA-DRsequences with known three-dimensional structure revealeddifferencesintheaminoacidcompositionofpocket9betweenDRB1*0401andDRB1*0405subtypes,andbetweenDRB1*0401andDRB1*0801allotypes.Onceagain,thisisconsistentwiththeresultingpocketprofiles(Fig.1C).Furthermore, a comparison of DRB1*0405 and DRB1*0801sequencesviaalignmenttothree-dimensionalstructuresindicated identical pocket 9 topographies and, consequently,predictedsimilarpocketprofiles.TheprofilesshowninFigure1Cdemonstratethatthiswasindeedthecase.We per-formedasimilarsetof experiments forthepolymorphicpocket6(datanotshown)andwereableto furtherconfirmthe approximationthatprofiles aremainly independentofthe remaining HLA-DR groove and that primary HLA-DRBstructuresaresufficienttoassignprofilestogivenHLA-DRpockets.

The approximation that pocket profiles show allele inde-pendence enables the generation of virtual matrices; that is,profilesforidenticalpocketsarerecycledfrom a poolrather

Table2.AssembledDRvirtualmatrices.

DRB1*0101[1;1;1;1;1;1]	DRB1*0102[2;1;1;1;1]	DRB1*1501[2;2;2;2;1]
DRB1*1502[1;2;2;2;1]	DRB1*1506[2;2;2;2;1]	DRB1*0301[2;3;3;3;2]
DRB1*0305[1;3;3;3;3]	DRB1*0306[2;3;3;4;3]	DRB1*0307[2;3;3;4;3]
DRB1*0308[2;3;3;4;3]	DRB1*0309[1;3;3;3;2]	DRB1*0311[2;3;3;4;3]
DRB1*0401[1;4;4;4;3]	DRB1*0402[2;5;4;5;3]	DRB1*0404[2;6;4;6;3]
DRB1*0405[1;6;4;6;5]	DRB1*0408[1;6;4;6;3]	DRB1*0410[2;6;4;6;5]
DRB1*0421[1;4;4;4;2]	DRB1*0423[2;6;4;6;3]	DRB1*0426[1;4;4;4;3]
DRB1*1101[1;7;3;7;3]	DRB1*1102[2;11;3;11;3]	DRB1*1104[2;7;3;7;3]
DRB1*1106[2;7;3;7;3]	DRB1*1107[2;3;3;3;3]	DRB1*1114[1;11;3;11;3]
DRB1*1120[1;11;3;11;2]	DRB1*1121[2;11;3;11;3]	DRB1*1128[1;7;3;7;2]
DRB1*1301[2;11;3;11;2]	DRB1*1302[1;11;3;11;2]	DRB1*1304[2;11;3;11;5]
DRB1*1305[1;7;3;7;2]	DRB1*1307[1;7;3;9;3]	DRB1*1311[2;7;3;7;3]
DRB1*1321[1;7;3;7;5]	DRB1*1322[2;11;3;11;3]	DRB1*1323[1;11;3;11;3]
DRB1*1327[2;11;3;11;2]	DRB1*1328[2;11;3;11;2]	DRB1*0701[1;8;5;8;4]
DRB1*0703[1;8;5;8;4]	DRB1*0801[1;9;3;9;5]	DRB1*0802[1;9;3;9;3]
DRB1*0804[2;9;3;9;3]	DRB1*0806[2;9;3;9;5]	DRB1*0813[1;9;3;6;3]
DRB1*0817[1;9;3;7;5]	DRB5*0101[1;10;6;10;6]	DRB5*0105[1;10;6;10;6]

Virtual DR matrices were assembled according to the modular structure of the HLA-DRgroove as indicated in Figure 1B. Profiles for pockets 4, 6, 7, and 9 were derived from thedatabaseshowninTable1.Profilesfortherelativepeptidepositions2and3werederivedfromtheDRB1*0401matrix[8](not shown).Forrelativepeptideposition1,onlyaliphatic(Ile,Leu,Met,Val)andaromatic(Phe,Trp,Tyr)aminoacidresidueswerec onsidered;morespecif-ically,forHLA-DRalleleswithaβ86Glycomposingpocket1,valuesof0wereassignedtoaromaticand-1toaliphaticresiduesatrelativeP1,whilethereversewasdoneforDRalleleswithaβ86Valcomposingpocket1(ref.8).Thevirtua lmatrixvaluesforeachallelearecod-ed by a set of five numbers listed after each allele: The first number indicates whether theallelehasaGly(=1)oraVal(=2)atβ86(seeabove).Thesecondnumberrepresentstheiden-tificationnumberofthepocket4profile(seeTable1).Thethird,fourth,andfifthnumberindi-catestheidentificationnumberofthepocket6,7,and9profiles,respectively.

than determined repeatedly for each allele. The important conse-quence is that a relatively small number of profiles can be used tobuildalargenumberofHLA-DRmatrices.Thesynthesisofapproxi-mately 1,000 synthetic designer peptides and the accomplishment of10,000HLA-ligandbindingassaysallowedustocreateadatabaseof35independentpocketprofiles(Table1).These35profileswereused to build 51 virtual HLA-DR matrices (Table 2), which represent themajorityofthehumanHLA-DRpeptidebindingspecificity[20].

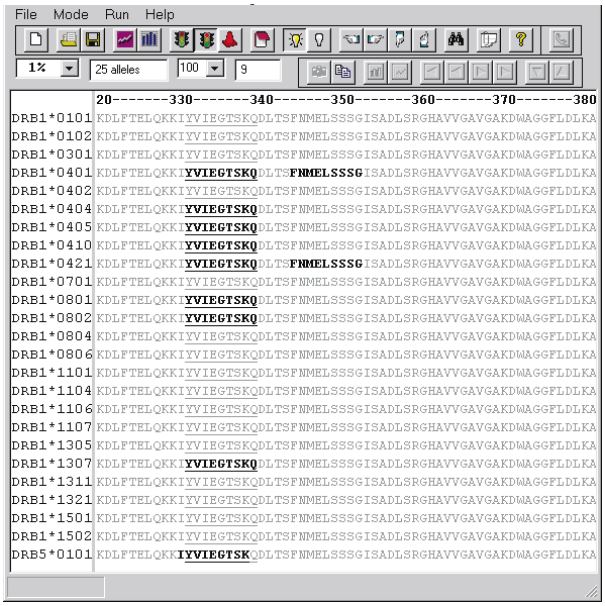
HLA-DR ligand prediction with virtual matrices. Are virtual matricessuitableforthepredictionofHLAclassIIligandsandcandi-date T-cell epitopes? To answer this question, we created a new soft-warepackagenamedTEPITOPE,inwhichthepocketprofilesandthe resulting virtual matrix data were incorporated (Fig. 2A). The basicligand prediction algorithm works, in principle, like earlier quantita-tive matrix-based algorithms (see above). However, instead of calcu-lating only peptide scores for every peptide frame in a given proteinsequence, it enables the calculation of score distribution curves foreachHLA-DRallotypeusingnaturalproteinsequencedatabasesasasource (Fig. 2B). Thus, peptides are predicted based on a user-select-ed threshold defined as the percentage of best scoring natural pep-tides (Fig. 2B). This compensates in part for the allelic differences ofabsolutepeptidescorescausedbyvariationsinthesensitivityofHLA-DRpeptidebindingassays(datanotshown).

The predictive power of virtual matrices was tested on both indi-vidual T-cell epitopes and large peptide repertoires. Gross et al. [18]haverecentlyusedourpreviouslydescribedquantitativeDRB1*0401matrixtoidentifyacandidateautoantigenicpeptidefo rLymearthri-tis. Figure 2A shows that the virtual matrices incorporated into oursoftware would have predicted this peptide too, using a stringentthreshold setting of “1% best scoring natural peptides.” Figure 2Cshows that our softward can also be used to determine ‘thresholdprofiles’forpeptides.Forexample,thethreshold profileofMAGE-3281–295, a peptide originally identified with TEPITOPE[21], revealedthat it is predicted to bind to many HLA-DR allotypes, even whenstringent threshold settings are used (Fig. 2C). Notably, we con-firmed the promiscuity of MAGE-3 281–295 by in vitro bindingstudies,andwealsodemonstratedthatMAGE-3281–295was indeedpresentedbymelanomacells[21].IncontrasttoMAGE-3281–295, the melanoma-specific helper T-cell epitope tyrosinase448–462 (ref. 22) was described as being a DRB1*0401-restrictedlow-affinity ligand. This again is consistent with the threshold pro-fileforthispeptide(Fig.2C).

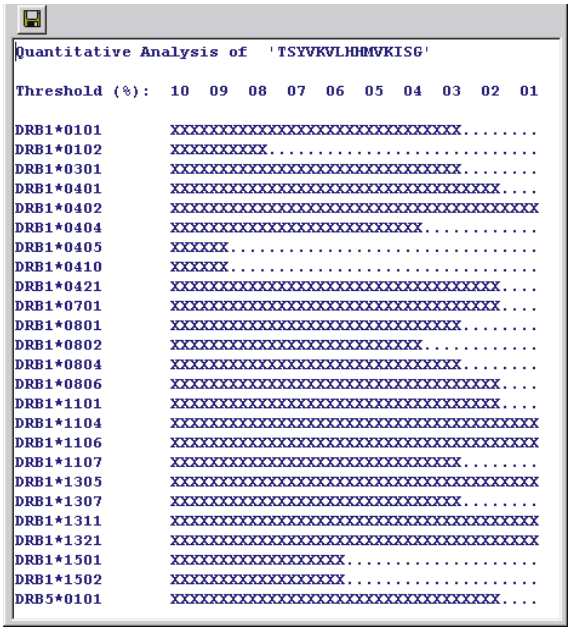
Obviously, larger ligand repertoires are required for a better esti-mation of the predictive power of virtual matrices. Therefore, wetested several molecular repertoires. The first repertoire consisted ofbothHLA-DR–selectedandnonselectedpeptidesoriginallygenerat-edbythebacteriophageM13displaytechnology[23,24].Wethentestedwhether we could computer-simulate the screening of M13 displaylibraries. We combined both the selected and nonselected peptiderepertoires and examined whether the virtual matrices could “sepa-rate” them again computationally (Fig. 3A). Up to 80% of the HLA-DR selected peptides could be predicted using a stringent thresholdsettingof1–3%,whereas<5%ofthenonselectedpeptideswerepre-dicted under the same conditions (Fig. 3B). These results clearlydemonstrated the ability of TEPITOPE to computationally separateHLA selected and nonselected peptide repertoires. To further assessthepredictivepowerofvirtualmatrices,weperformedpeptidebind-ingassays withhundredsofrandomly selectednaturalpeptidesequences, generating yet another repertoire of HLA-DR bindingand nonbinding peptides. We demonstrated that stringent thresholdsettings were sufficient for the preferential prediction of HLA-DRligands (data not shown). Finally, we tested natural ligands and T-cellepitopesusingthenaturalliganddatabasegeneratedbyRammensee’sgroup[25].Morethanhalfofallnaturalligandscouldb epredicted using a 1–3% threshold setting and more than 75% with a1–6%thresholdsetting(Fig.3C).Inconclusion,theuseoflargedatasets that were either derived experimentally in our laboratory (Fig.3B and data not shown) or from the literature (Fig. 3C) demonstrat-ed the utility of the threshold setting for the prediction of HLA-DRligands.Inaddition,itallowedustoestimatethepotentialfalse-pos-itive and false-negative rate at different threshold stringencies (Fig.3Banddatanotshown).

Generation of promiscuous HLA-DR ligand databases. ThecomputationalpredictionofcandidateT-cellepitopesbyvirtual

matrices is not limited to well-defined protein sequences. Variousgenome projects are generating huge amounts of new sequenceinformation[26–29],andhigh-throughputsequencingofcDNAlibraries has led to the discovery of several millions of expressedsequencetags(ESTs)[30,31].Theavailabilityofthesesequencesmakesit possible to quantify mRNA levels for tens of thousands of genessimultaneously byusinghigh-density oligonucleotidearrays[16,17].Moreover,comparativetranscriptprofilingstudieswithDNAmicroarraysenablethediscovery oflargerepertoiresofgenesthat

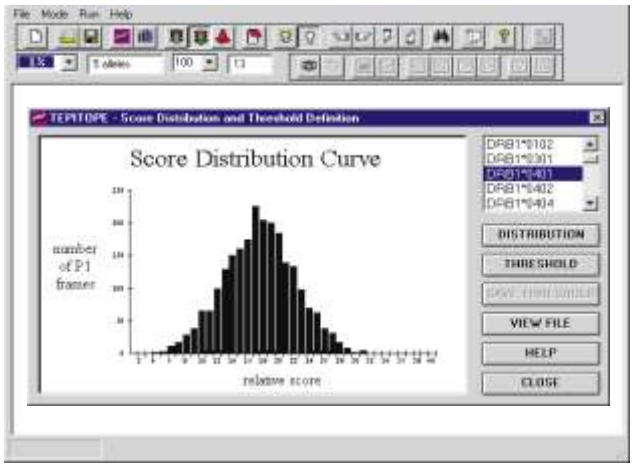


C



areeitherspecificallyexpressedorupregulatedindiseasetissues(datanotshown). WeproposetoemployTEPITOPEonagenome-widelevelforthegenerationofcomprehensiveHLA-DRliganddatabases.Forexam-ple,helperTcellshavebeenshowntoplayacrucialroleftheopti-malinductionofprotectiveimmunityagainstcertaintypesoftumors[32]. A database ofpromiscuous candidateT-cell epitopesofgenesupregulatedorspecificallyexpressedintumortissuescouldbeavaluabletoolforthedesignofepitope-basedvaccines.Todemon-

B




```
Quantitative Analysis of  'DYSYLQDSDPDSFQD'

Threshold (%):  10  09  08  07  06  05  04  03  02  01

DRB1*0101       ..................................
DRB1*0102       ..................................
DRB1*0301       XXXXXX............................
DRB1*0401       XXXXXXXXXXXXXXXXXXXXXXXXXX........
DRB1*0402       ..................................
DRB1*0404       ..................................
DRB1*0405       ..................................
DRB1*0410       ..................................
DRB1*0421       XXXXXXXXXXXXXXXXXXXXXX............
DRB1*0701       ..................................
DRB1*0801       ..................................
DRB1*0802       ..................................
DRB1*0804       ..................................
DRB1*0806       ..................................
DRB1*1101       ..................................
DRB1*1104       ..................................
DRB1*1106       ..................................
DRB1*1107       XXXXXXXXXXXXXX....................
DRB1*1305       ..................................
DRB1*1307       ..................................
DRB1*1311       ..................................
DRB1*1321       ..................................
DRB1*1501       ..................................
DRB1*1502       ..................................
DRB5*0101       ..................................
```

Figure 2. Function of the TEPITOPE software. (A) User interface and prediction of a selective peptide in human leukocyte function–associatedantigen-1. The predicted region (bold) corresponds to a recently described candidate autoantigenic peptide (underlined) for Lyme arthritis(humanleukocytefunction–associatedantigen-1,hLFA-1□L332–340,[ref.18]).Thepredictionthresholdwassetto1%(Fig.2B).(B)Calculationanddisplayofscoredistributioncurves.TEPITOPEallowsthecalculationofscoredistributioncurvesforeachHLA-DRallelebasedonanynatural protein database. The Figure shows the DRB1*0401 score distribution of all possible peptide frames in a database of natural proteinsequences (8,000 peptide frames). This database is used to normalize the prediction for each HLA-DR allele. Prediction thresholds (chosen bythe operator) are expressed as percentage of the best scoring peptides in natural peptide frames. (C) Quantitative evaluation o f thresholdprofiles for given peptides. For **any submitted peptide sequence, a**

histogram displays the predictability for each DR allele according to thethreshold stringency: bars indicate the threshold setting at which the peptide is predicted as a ligand for each listed DR alle le. Examples ofquantitative evaluations are shown for DR promiscuous MAGE-3 281–295[21] (left) and allele-specific DRB1*0401 restricted (tyrosinase 448–462[ref.22])(right)peptidesderivedfromtumor-associatedantigens.

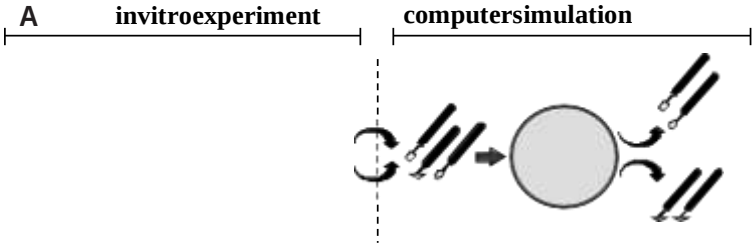
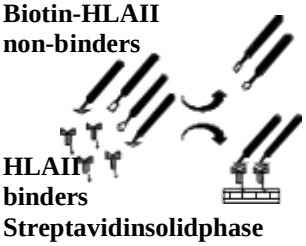


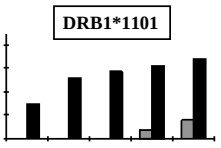
Figure3.ValidationofTEPITOPEusinglargepeptiderepertoires.(AandB)Simulationofabacteriophage

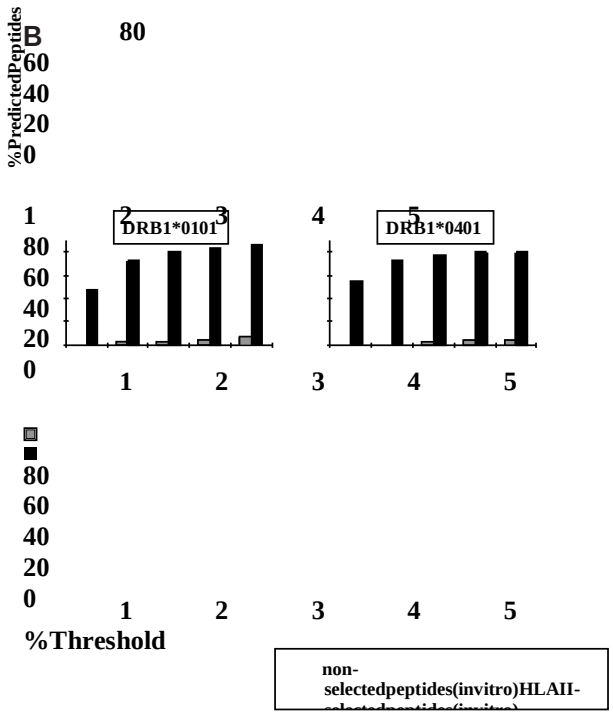
Phagepeptidedisplaylibrary



Predictionsoftware
non-
binders

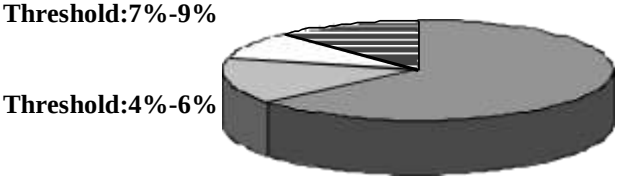
HLAII
binders
peptidedisplaylibraryscreeningwithHLA-DRmolecules.(A)Schematicrepresentationofthevalidation experiment: (left) bacteriophage-displayedpeptides binding to biotinylated DR molecules areseparatedbysolidphase-boundstreptavidincapture,and isolated by low pH elution; (right) the peptidepoolsgeneratedwiththeM13bacteriophagedisplaytechnologyaresubmittedtoanalysisforthecomputational"isolation"ofDR-selectedpeptides.
(B) Prediction of in vitro generated DR repertoiresconsisting of 60 DRB*0101-selected, 52 DRB*0401-selected,52DRB*1101-selected,and60nonselectedbacteriophagedisplayedpeptidesequences.Results foreachHLA-DRalleleareexpressedasthe





percentage of predicted peptides out of the total ofthe respective HLA-DR-selected (black bars) andnonselected (gray bars) peptide repertoires at theindicated threshold settings. (C) Prediction of HLA-DRbindersfromthenaturalHLA-DRpeptideligandsandT-cellepitopesdatabase[25].NaturalHLA-DRpeptide ligands and T-cell epitopes (n = 223)wereselected for prediction based on the following twocriteria:(1)asmanypeptidesarerepresentedassetsoftruncatedpeptides,onlyonesequencepersetwasselected,and(2)asthe currentalphaversionof

HLAIIT-cellepitopeandnaturalliganddatabase



TEPITOPE represents only 25 alleles, only availableligandsforthesealleleswereselected(Fig.2A).Thepeptides were analyzed at the indicated differentthreshold settings. Results are expressed for all HLA-DR alleles as the percentage of predicted peptidesout of the total of the peptide repertoire. Predictedpeptidesareonlycountedonceforagiven HLA-DRallele; that is, a peptide predicted at a stringentthreshold(e.g.,1–3%)wouldnotbecountedagainatalessstringentthresholdsetting(e.g.,4–6%).

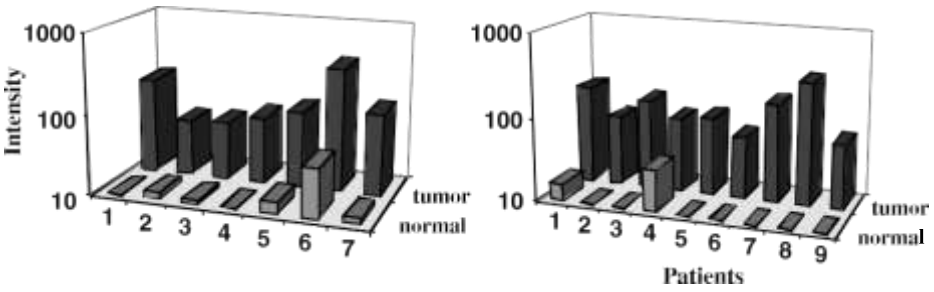


Figure4.Examplesfortumorantigenidentificationby DNAmicroarraytechnology.C-mycandanendogenousretroviralproteasewereupregulatedin7/20and9/20coloncancerpatients,respectively.Antibodiesh avebeendescribed in the serum of cancer patientsforbothantigens[36](datanotshown).Quantification for any mRNA is given bythe sum of all perfect match intensitiessubtractedfrommismatchintensitiesdivided by the total number of probe pairs(='intensity')[16].

stratethatsuchacomprehensivedatabasecaneasilybegeneratedbycombining DNA microarraytechnology withepitopepredictionsoftware, we performed a simple pilot study: Using both a commer-cially available Affymetrix (Santa Clara, CA) DNA microarray set(~7,000 genes) and two of our own microarray designs (~12,000genes; Fig. 4A and B), we have recently profiled 20 primary coloncancer tissues together with the corresponding adjacent normal tis-sues(datanotshown). Althoughmorethan1,000independentgeneswerefoundtobedifferentiallyexpressedinapopulationof20patients,only34ge neswereupregulatedorspecificallyexpressed in

 50% of all patients (data not shown). These 34 genes gave rise toapproximately19,000peptideframes.Ofthese19,000peptideframes, 130 candidate promiscuous T-cell epitopes were predictedby TEPITOPE using the following criteria: First, threshold (1–3%best scoring natural peptides); second, promiscuity (predicted tobind to 5/7 HLA-DR allotypes); and third, peptide length (15 aminoacid residues). This example demonstrates both the relative ease ofgeneratingsuchadatabaseandthemanageabledataoutput.

Moreover, the fact that antibodies have been described in serum ofcancer patients for some of the microarray-selected candidate anti-gens(Fig.4)furthersupportsthefeasibilityofsuchanapproach.

DatabasesofcandidateHLA-DRligandsandhelperT-cellepitopescouldultimatelybedeterminedforeverygeneinagenome.However,the combination of epitope prediction software with other "filters," asdemonstratedinthisreport,willobviouslybemorepractical.DNAmicroarray/predictionsoftware–basedapproachestogeneratedata-basesofpromiscuouscandidateT-cellepitopescouldbewidelyapplicable in other areas. For example, the current genome project forthemalaria-causingpathogenPlasmodiumfalciparumshouldsoonmake it possible to generate similar databases (e.g., for life cycle–spe-cific candidate antigens). Similarly, approaches that use epitope pre-dictionsoftwareincombinationwithserologicalidentificationofantigensbyrecombinantexpressioncloning(SEREX)tech nology[33]might alsoproveveryuseful.SEREXallowsthesystematicidentifica-tionofantigensinhumancancersandhasledtothedefinitionofawealthofnewtumorantigensinmanydifferenttumorentities.

Experimentalprotocol

Determinationofapocketprofiledatabase. Pocketprofileswerederived nucleotides[16].HybridizationofcRNAtothemicroarraysandquantificationofRNAexpressionwasperformedasdescribedin ref.16.

fromside-chainscanningdataobtainedbysubstitutingallele-specificpep-

tide ligands (basis peptides) in position 4, 6, 7, and 9, with all natural aminoacid residues. Peptide interactions with detergent-solubilized HLA-DR mol-ecules weremeasuredusing an ELISA-basedhigh-throughput competitivebindingassayas described[13,34].ForeachHLA-DRmoleculeanalyzeda spe-cific basispeptide was selectedafterseveraloptimizationexperimentstoguarantee a highly sensitive analysis of the effects of each peptide side chainon HLA-DR binding[13]. The following basis peptides were used in this study:Gly-Phe-Lys-Ala-Ala-Ala-Ala-Ala-Ala-Ala forDRB1*0101,DRB5*0101,and DRB1*0701; Ile-Ala-Tyr-Asp-Ala-Ala-Ala-Ala-Alafor DRB1*0301; Tyr-Arg-Ser-Met-Ala-Ala-Ala-Ala-Alafor DR1*0401,DRB1*0801,andDRB1*1101;Gly-Ile-Arg-Ala-Ala-Tyr-Ala-Ala-Ala-AlaforDRB1*1501.Competitionassays wereconductedtomeasuretheability of substitutedbasispeptidesto competewithabiotinylatedindicatorpeptidesforbindingtopurifiedDRmolecules. Atleastfivedilutionsweredeterminedforeachcompetitorpeptide. The resulting data points were plotted [34]andthe shapeof the curves were used for quality control; that is, data sets that did not dis-play asigmoidshape wererepeated.The following biotinylatedindicatorpeptides wereused:Gly-Phe-Lys-Ala-Ala-Ala-Ala-Ala-Ala-Ala-AlaforDRB1*0101andDRB1*0701,Gly-Ile-Arg-Ala-Ala-Tyr-Ala-Ala-Ala-AlaforDRB1*1501, myelin-based protein 85–99 for DRB5*0101, Ile-Ala-Tyr-Asp-Ala-Ala-Ala-Ala-AlaforDRB1*0301,Tyr-Pro-Lys-Phe-Val-Lys-Gln-Asn-Thr-Leu-Lys-Ala-AlaforDRB1*0401(ref.19),tetanustoxoid$_{830–843}$for DRB1*1101(ref.35),andGly-Tyr-Arg-Ala-Ala-Ala-Ala-Ala-Ala-LeuforDRB1*0801.Therelativebindingdataofthecompetitorpeptides wereexpressedasthe concentrationof competitorpeptiderequiredtoinhibit50%ofbindingofthebiotinylatedindicatorpeptide(IC_{50}).

Assembly of virtual matrices and software. Virtualmatrices were assem-bledas follows.First,multiplealignments ofHLA-DR sequencestoDRsequencesof knownthree-dimensionalstructures wereperformedtolinkpolymorphic DR residues to given DR pockets. Second, pockets were classi-fied according to their composition of polymorphic residues; that is, pocketsfrom different alleles constituted by identical residues were considered iden-tical.Third,apocketprofilesdatabasewasdeterminedinvitroon11HLA-DR alleles(Table 1).Fourth,pocketprofileswereassignedtopockets of allHLA-DRallelesaccordingtotheirclassification.Andfifth,51fullyassembledvirtualDRmatrices were generatedby combiningthe assignedprofiles ofpockets4,6,7,and9,andcleftregion2and3.Expertruleswereusedforpocket1,aspreviouslydescribed[19](Table2,legend).P rofilesforpeptideposi-

tions5and8werenotconsideredduetotheirminimaleffectonbinding[3](datanotshown).TEPITOPEisaWindows98/NTappli cation.Thevisualuserinterfaceallowstheidentificationofpromiscuous HLA-DR ligandsindependentof whetheridenticaland/orshiftedHLA-DR binding framesconstitute promiscuity. Twenty-five virtual HLA-DR matrices were incorpo-rated intothe currentalpha versionof the application. Requests touse thealpha version of TEPITOPE should be addressed via e-mail to juergen.ham-mer@roche.com.

Microarraydesign,RNAsamplepreparation,hybridization,andanalysis. Three microarray designs were used for transcript mapping of primary coloncancertissue.Microarray1iscommerciallyavailable(6.8kHumanChip;Affymetrix,SantaClara,CA)andcovers7,071 genes.Microarrays2and3arecustomdesigns,eachcovering6,088genes.Thecommerciallyavailablemicroarray 1 represents the currently known set of functionally characterizedgenes,whichareallavailableinthepublicdomain.Thedesignofmicroarrays2 and 3 will be described elsewhere (data not shown). In brief, microarray 2consists oftranscriptsforwhichhigh-qualityconsensussequencescouldbegenerated from public and proprietary EST databases. "High quality" meansthat the consensus is based on an EST sequence redundancy of at least five tocorrect for the vast majority of EST sequencing errors. The minimum lengthrequirement of 500 nucleotides applied in this design is significantly exceed-edbymostofthesequences.Microarray3contains3□ sequenceswithamini-mum of five contributing ESTs from public and proprietary sources. In addi-tion, sequences were selected to exclude any significant sequence homologybetweengenesrepresentedonthemicroarrayset.

RNA was extractedfromprimary coloncancerandadjacentnormalhumantissueusingtheUltraspecmethod

(Biotecx,Houston,TX).RNAwasconvertedintocDNAbyreversetranscriptionandthenintocRNAwithaninvitrotranscription reactionthatcontainedbiotin-labeledCTPand UTP

REFERENCE

1. Germain, R.N. MHC-dependent antigen processing and peptide presentation:providingligandsforTlymphocyteactivation.Cell**76,**287–299(1994).
2. Topalian, S.L. MHC class II restricted tumor antigens and the role of CD4 T cellsincancerimmunotherapy.Curr.Opin.Immunol.**6,**741–745(1994).
3. Stern, L.J. et al. Crystal structure of the human class II MHC protein HLA-DR1complexedwithanInfluenzaviruspeptide.Nature**368,**215–221(1994).
4. Brown, J.H. et al. The three-dimensional structure of the human class II histo-compatibilityantigenHLA-DR1.Nature**364,**33–39(1993).
5. Ghosh,P.,Amaya,M.,Mellins,E.&Wiley,D.C.Thestructureofanintermediateinclass IIMHCmaturation:CLIPboundtoHLA-DR3.Nature**378,**457–462(1995).
6. Dessen, A., Lawrence, M.C., Cupo, S., Zaller, D.M. & Wiley, D.C. X-ray crystalstructure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide fromhumancollageneII.Immunity**7,**473–481(1997).
7. Sinigaglia, F. & Hammer, J. Defining rules for the peptide-MHC class II interac-tion.Curr.Opin.Immunol.**6,**52–56(1994).
8. Hammer, J. et al. Precise prediction of major histocompatibility complex class IIpeptide interaction based on peptide side chain scanning. J. Exp. Med. **180,**2353–2358(1994).
9. Marshall,K.W.etal.PredictionofpeptideaffinitytoHLADRB1*0401.J.Immunol.**154,**5927–5933(1995).
10. Brusic,V.,Rudy,G.,Honeyman,G.,Hammer,J.&Harrison,L.PredictionofMHCclass II-binding peptides using an evolutionary algorithm and artificial neural net-work.Bioinformatics**14,**121–30(1998).
11. Sette, A. et al. Structural characteristics of an antigen required for its interactionwithIaandrecognitionbyTcells.Nature**328,**395–399(1987).
12. Hammer, J., Sturniolo, T. &Sinigaglia, F. HLA class II peptide binding specificityandautoimmunity.AdvImmunol**66**,67-100(1997).
13. Hammer, J. &Sinigaglia, F. in MHC Vol. 2, Vol. 181 (eds Fernandez, N. & Butcher,G.)197–228(OxfordUniv.Press,NewYork,1998).
14. Hammer, J. New methods to predict MHC-binding sequences within protein anti-gens.Curr.Opin.Immunol.**7,**263–269(1995).
15. Marsh, S.G.E. Nomenclature for factors of the HLA system, update January1998.TissueAntigens**51,**582–583(1998).
16. Lockhart,D.J.etal.Expressionmonitoringbyhybridizationtohigh-densityoligonucleotidearrays[seecomments].Nat.Biotechnol.**14,**1675–1680(1996).
17. Wodicka, L., Dong, H., Mittmann, M., Ho, M.H. & Lockhart, D.J. Genome-wideexpressionmonitoringinSaccharomycescerevisiae.Nat.Biotechnol.**15,**1359–1367(1997).
18. Gross,D.M.etal.IdentificationofLFA-1asacandidateautoantigenintreatment-resistantLymearthritis.Science**281,**703–706(1998).
19. Hammer, J. et al. Peptide binding specificity of HLA-DR4 molecules: correlationwithrheumatoidarthritisassociation.J.Exp.Med.**181,**1847–1855(1995).
20. Tsuji,K.,Aizawa,M.&Sasazuki,T.HLA1991;Proceedingsoftheeleventhinter-national histocompatibility workshop and conference (Oxford Univ. Press, NewYork,1992).
21. Manici, S. et al. Melanoma cells present a MAGE-3 epitope to CD4(+) cytotoxic Tcells in association with histocompatibility leukocyte antigen DR11. J. Exp. Med.**189,**871–876(1999).
22. Topalian, S.L. et al. Melanoma-specific CD4 T cells recognize nonmutated HLA-DR-restrictedtyrosinaseepitopes.J.Exp.Med.**183,**1965–1971(1996).
23. Hammer,J.,Takacs,B.&Sinigaglia,F.IdentificationofamotifforHLA-DR1bind-ingpeptidesusingM13displaylibraries.J.Exp.Med.**176,**1007–1013(1992).
24. Hammer, J. et al. Promiscuous and allele-specific anchors in HLA-DR bindingpeptides.Cell**74,**197–203(1993).
25. Rammensee,H.G.,Bachmann,J.,Emmerich,N.&Stevanovic,S.SYFPEITHI:aninternet database for MHC ligands and peptide motifs (access via http:/www.uni-tuebingen.de/uni/kxi/).(dataextractedin4Q,1998).
26. Blackwell, J.M. Parasite genome analysis. Progress in the Leishmania genomeproject.Trans.R.Soc.Trop.Med.Hyg.**91,**107–110(1997).
27. Collins,F.S.etal.NewgoalsfortheU.S.HumanGenomeProject:1998-2003.Science**282,**682–689(1998).
28. Harwood, C.R. &Wipat, A. Sequencing and functional analysis of the genome ofBacillussubtilisstrain168.FEBSLett.**389,**84–87(1996).
29. Degrave,W.,Levin,M.J.,daSilveira,J.F.&Morel,C.M.Parasitegenomeprojectsand the Trypanosomacruzi genome initiative.Mem. Inst. Oswaldo Cruz **92**,859–862(1997).
30. Adams, M.D. et al. Sequence identification of 2,375 human brain genes [seecomments].Nature**355,**632–634(1992).
31. Adams, M.D. et al. Complementary DNA sequencing: expressed sequence tagsandhumangenomeproject.Science**252,**1651–1656(1991).
32. Ossendorp,F.,Mengede,E.,Camps,M.,Filius,R.&Melief,C.J.SpecificThelpercellrequirementforoptimalinductionofcytotoxicTlymphocytesagainstmajorhis-tocompatibilitycomplexclassIInegativetumors.J.Exp.Med.**187,**693–702(1998).
33. Sahin, U., Türeci, Ö. &Pfreundschuh, M. Serological identification of humantumorantigens.Curr.Opin.Immunol.**9,**709–716(1997).
34. Raddrizzani, L. et al. Different modes of peptide interaction enable HLA-DQ andHLA-DRmoleculestobinddiversepeptiderepertoires.J.Immunol.**159,**703–711(1997).
35. Panina-Bordignon, P. et al. Universally immunogenic T cell epitopes: promiscu-ous binding to human MHC class II and promiscuous recognition by T cells. Eur.J.Immunol.**19,**2237–2242(1989).
36. Ben-Mahrez, K., Thierry, D., Sorokine, I., Danna-Muller, A. &Kohiyama, M.Detection of circulating antibodies against c-myc protein in cancer patient sera.Br.J.Cancer**57,**529–534(1988).