tions. Thus the difference between the region around Tyr^{90} in the NMR model and the crystallographic studies reflects a genuine difference between the solution and the crystal conformation. The differences around the loop encompassing residues 31 to 36, however, although arising from crystal interactions, are misleading, because as shown by the joint x-ray–NMR refinement and the concomitant lack of NOE violations, the crystallographic model is consistent with the NMR observations (*15*).

An attractive aspect of the joint x-ray–NMR refinement method is a decreased necessity for human intervention in rebuilding the model throughout the process. Because the NMR observations ensure maintenance of proper local interactions between side chains, they can assist in selecting side chain conformations during interpretation of the electron density map, as demonstrated for example in the case of Leu^{69} and Leu^{82}. Several NOE interactions between the side chain protons of these two residues and protons of the adjacent Leu^{26} and Leu^{80}, as well as dihedral NMR restraints on the χ_1 angles, lock their side chains in a conformation that is different from the one observed in the refinements without NOE restraints (models 4I1B and NONOE). The resulting electron density map confirms the choice of conformation for both Leu^{69} and Leu^{82} (Fig. 2). Unlike the maps produced by the refinements without NOEs (Fig. 2B), it also indicates that alternative conformations are possible, particularly for Leu^{82} (Fig. 2A). Such amplification of the information contained in the diffraction data, together with the improvement in the ratio of parameters to observations, suggest that incorporation of the NMR observations, where available, has the potential of improving the convergence of crystallographic refinement procedures.

Although previous studies have shown that structures determined by NMR can be used as initial models for solving crystal structures by the molecular replacement method (*16*), the present study opens the way for a full amalgamation of the NMR and x-ray data throughout the structure determination process. A typical situation in which the diffraction data can be complemented by the NMR observations is in studies of multidomain proteins, for which only the overall structure can be modeled reliably from medium to low-resolution electron density maps, but more detailed models of the individual domains can be obtained from NMR data pertaining to higher resolution.

REFERENCES AND NOTES

1. K. Wüthrich, *Science* **243**, 45 (1989); G. M. Clore and A. M. Gronenborn, *ibid.* **252**, 1390 (1991).
2. M. Billeter *et al.*, *J. Mol. Biol.* **206**, 677 (1989); T. A. Holak *et al.*, *ibid.* **210**, 649 (1989); G. M. Clore and A. M. Gronenborn, *ibid.* **217**, 611 (1991); in *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy*, J. C. Hoch *et al.*, Eds. (Plenum, New York, 1991), pp. 57–65.
3. G. M. Clore and A. M. Gronenborn, *J. Mol. Biol.* **221**, 47 (1991).
4. A. Jack and M. Levitt, *Acta Crystallogr.* **A34**, 931 (1978).
5. W. A. Hendrickson and J. H. Konnert, in *Computing in Crystallography*, R. Diamond, S. Ramaseshan, K. Venkatesan, Eds. (Indian Academy of Sciences, Bangalore, 1980), pp. 13.01–13.23.
6. A. T. Brünger, G. M. Clore, A. M. Gronenborn, M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3801 (1986); A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987); A. T. Brünger, G. M. Clore, A. M. Gronenborn, M. Karplus, *Protein Eng.* **1**, 399 (1987); A. T. Brünger, *X-PLOR, Version 2.1* (Yale University, New Haven, CT, 1990).
7. G. M. Clore, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* **30**, 2316 (1991).
8. J. P. Priestle, H. P. Schär, M. G. Grütter, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9667 (1989).
9. B. C. Finzel *et al.*, *J. Mol. Biol.* **209**, 779 (1989).
10. B. Veerapandian *et al.*, *Proteins* **12**, 10 (1992).
11. The data used for the joint refinement comprised NMR observations, consisting of 2,780 experimental distance restraints and 390 experimental dihedral restraints (*7*) and 10,429 x-ray reflections extending to 2.0 Å resolution (*10*). Three rounds of simulated annealing with slow cooling (*6*), followed by minimization and restrained temperature factor refinement, which were performed at resolution shells of 2.8, 2.5, and 2.0 Å, caused the *R* value [$R = \Sigma_h|F_0(\mathbf{h}) - F_c(\mathbf{h})|/\Sigma_h\mathbf{F}_0(\mathbf{h})$, where $F_0(\mathbf{h})$ and $F_c(\mathbf{h})$ are the observed and calculated structure factors, respectively] to decrease from the initial value of 0.45 to 0.228 for all of the data to 2.0 Å. Finally, several rounds of minimization at 2.0 Å resolution with the more restrictive PROLSQ (*5*)–like geometry parameters brought the *R* value down to 0.214. Only the six water molecules for which NMR observations could be made (*7*) were included in the model throughout the refinement. In order to monitor the effect of combining the NMR observations with the x-ray data, a conventional crystallographic simulated annealing refinement, starting from the NMR model 6I1B, was performed in parallel to the joint refinement and along the same scheme, yielding an *R* value of 0.216 (model NONOE). In order to calculate the free *R* value of the joint refinement (JR) and the 4I1B models, a cycle of simulated annealing at 600 K was performed, leaving out 10% of the reflections selected at random as described in (*13*). The free *R* values obtained were 0.254 and 0.286 for the JR and 4I1B models, respectively.
12. M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* **229**, 317 (1988).
13. A. T. Brünger, *Nature* **355**, 472 (1992).
14. A. Wlodawer and W. A. Hendrickson, *Acta Crystallogr.* **A38**, 239 (1982).
15. The observation that the 31 to 36 loop in the ensemble of NMR structures does not overlap with the x-ray structures and that this difference stems from a rigid-body hinge movement about residues 31 and 36, indicates that the position adopted in the ensemble of NMR structures probably arise from the nonbonded potential term in the target function (in this case solely a quartic van der Waals repulsion term).
16. A. T. Brünger *et al.*, *Science* **235**, 1049 (1987); W. Braun, O. Epp, K. Wuthrich, R. Huber, *J. Mol. Biol.* **206**, 669 (1989); E. T. Baldwin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 502 (1991).
17. C. Chothia, *J. Mol. Biol.* **105**, 1 (1976).
18. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
19. T. N. Bhat and G. H. Cohen, *J. Appl. Crystallogr.* **17**, 244 (1984).

20. Supported by The AIDS Targeted Anti-Viral Program of the Office of the Director of National Institutes of Health (G.M.C., D.R.D., and A.M.G.). The coordinates of the IL-1β model derived by the joint x-ray–NMR refinement will be deposited in the Brookhaven Protein Data Bank.

5 May 1992; accepted 23 June 1992

A Critical Role for Conserved Residues in the Cleft of HLA-A2 in Presentation of a Nonapeptide to T Cells

France Latron, Laszlo Pazmany, Joanna Morrison, Robert Moots, Mark A. Saper,* Andrew McMichael, Jack L. Strominger

The peptide binding cleft of the class I human histocompatibility antigen, HLA-A2, contains conserved amino acid residues clustered in the two ends of the cleft in pockets A and F as well as polymorphic residues. The function of two conserved tyrosines in the A pocket was investigated by mutating them to phenylalanines and of a conserved tyrosine and threonine in the F pocket by mutating them to phenylalanine and valine, respectively. Presentation of influenza virus peptides and of intact virus to cytolytic T lymphocytes (CTLs) was then examined. The magnitude of the reduction seen by the mutation of the two tyrosines in the A pocket suggests that hydrogen bonds involving them have a critical function in the binding of the NH_2-terminal NH_3^+ of the peptide nonamer and possibly of all bound peptide nonamers. In contrast, the mutations in the F pocket had no effect on CTL recognition.

The class I and class II glycoproteins encoded in the major histocompatibility complex (MHC) of all vertebrates examined have a central role in immune recognition; they present peptides derived from foreign antigens to the T cell receptor complex on the effector cells of the immune system. These molecules are extremely polymorphic, which leads to the rejection of allogeneic or xenogeneic grafts. The elucidation of the crystal structures of two class I human molecules, HLA-A2 and HLA-

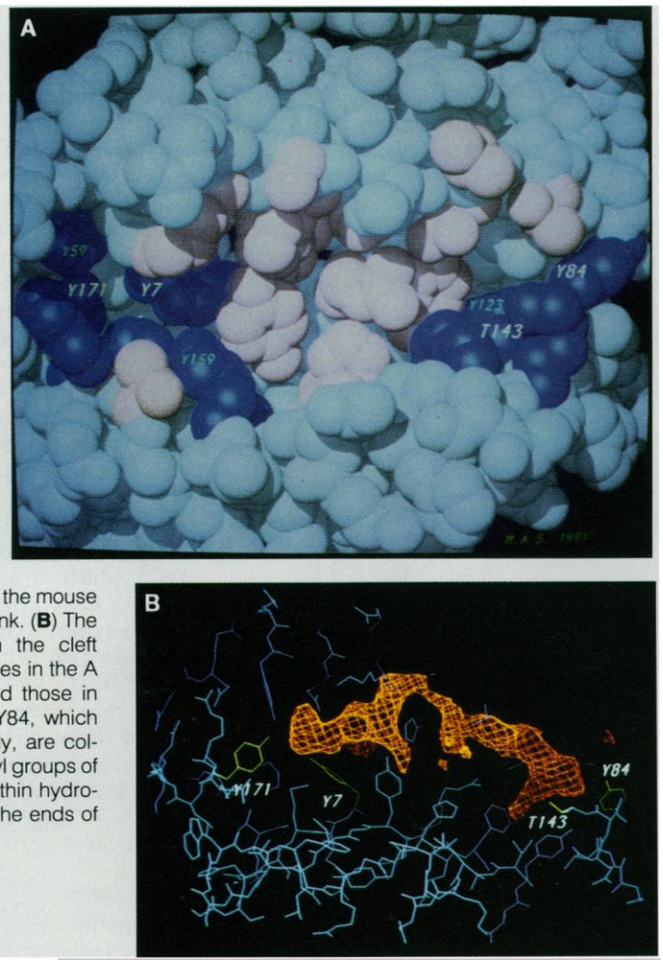

Fig. 1. The peptide binding site of HLA-A2.1 (*1, 2, 4*). (**A**) Space filling model of conserved residues in pockets A and F are colored dark blue, and polymorphic residues pink. The four tyrosine residues in pocket A (left) are numbered with the two residues mutated in this study, Y7 and Y171, highlighted. The conserved threonine and two tyrosine residues in the F pocket (right) are similarly numbered with Y84 and T143 highlighted. The tyrosine residue Y99 to the right of and between Y7 and Y159 is conserved in man although it is polymorphic in the mouse and is therefore colored pink. (**B**) The extra electron density in the cleft (shown in red). The residues in the A pocket, Y171 and Y7, and those in the F pocket, T143 and Y84, which were mutated in this study, are colored in green. The hydroxyl groups of these four residues are within hydrogen bonding distance of the ends of the extra electron density.

Aw68, and more recently a third, HLA-B27, revealed that self peptides are bound in a cleft in these molecules within their α1 and α2 domains (*1–5*). The sides of these clefts are two α helices (one from each domain), and their floor is those portions of a platform of eight β strands (four from each domain) that pass beneath the helices. All of the known polymorphic residues that characterize these molecules are located in or on this cleft. Residues on the walls and the floor of the cleft are presumed to be peptide contact residues, whereas residues on the tops of the helices are probably involved in recognition by T cell receptors. In previous studies the function of polymorphic residues in recognition by antibody, and by both virus specific and allogeneic cytotoxic T lymphocytes, has been investigated by mutagenizing specific amino acid residues (*6–8*). We have mutagenized some of the conserved residues in the peptide binding cleft and found that several of them function in binding the conserved feature of all peptides, their amino and carboxyl ends. In these studies we used a peptide nonamer that sensitizes HLA-A2 target cells for CTL lysis at 10^{-10} M.

The clefts of HLA-A2, -Aw68, and -B27 contain a number of pockets or depressions, called pockets A to F in HLA-A2, that are presumed to be the sites for binding of amino acid side chains (*3–5*). The nature of these pockets thus determines the peptide specificity of a given allele, and correspondingly, pockets B, C, D, and E in HLA-A2 contain polymorphic amino acid residues. The peptide binding cleft also contains conserved residues, clustered at the two ends of the cleft in pockets A and F (Fig. 1). Pocket A includes the conserved tyrosine (Y) residues Y7, Y59, Y159, and Y171 (whose hydroxyls all point into the pocket), whereas pocket F contains conserved tyrosines Y84 and Y123, Thr^{143} (whose hydroxyls are also within the pocket), and Lys^{146}, whose positive charge overhangs this pocket. In an atomic model of how a mixture of peptides binds to HLA-B27 (*5, 9*), the NH_2-terminal amino acid (including its $-NH_3^+$) of an extended nonameric peptide is bound in the A pocket, whereas its COOH-terminal amino acid (including its $-COO^-$) is bound in the F pocket through hydrogen bonds involving the two groups of conserved residues. The compensating mutations, $Tyr^{116} \rightarrow Asp^{116}$ in HLA-A2 and $Leu^{66} \rightarrow Arg^{66}$ in the peptide epitope, had also led to the conclusion that the COOH-terminus of the peptide is in the F pocket (*10*). We have mutagenized several of the conserved residues in the two pockets to examine the functional importance of the proposed hydrogen bonds that appear to involve them.

T cell epitopes that bind to class I MHC molecules are mainly peptide nonamers (*11–13*). In the case of HLA-A2, a mixture of peptide nonamers was found in the peptide binding cleft with the hydrophobic amino acids Leu and Met, and to a lesser extent Ile, predominating at the second position (P2) and Val or Leu at P9 (referred to as anchor positions) (*12*). Amino acids 57 to 68 of the influenza type A matrix protein (M57–68) are a T cell epitope (*14*), and M58–66 of this epitope (GlyIleLeuGlyPheValPheThrLeu) is the peptide nonamer with the appropriate anchor residues (*15, 16*). This peptide titrated out at about 10^{-10} M in the CTL assay. When the adjacent nonamers, M57–65, M59–67, and M60–68, are used, no lysis occurs, although M59–68 is recognized by some CTL clones at 10^{-6} M (*17*). When the long peptides M58–68 or M57–68 (LysGlyIleLeuGlyPheValPheThrLeuThrVal) are used, CTL lysis titrates out at about 10^{-8} M—about a 2-log reduction in effectiveness (*16*). We used HLA-A2 in which Tyr^{171} had been mutated to phenylalanine (Y171F) and was transfected into C1R cells. Recognition of M58–66 was reduced about 2 logs; recognition of M57–68 was also reduced, although to a smaller degree (Fig. 2, A and B). Recognition of the mutant in which Tyr^{7} was mutated to Phe (Y7F) was reduced in a manner similar to that of Y171F. These results were consistent with an observed ~50% reduction of presentation of the natural epitope (presumably M58–66) formed after virus infection with both Y171F and Y7F (Fig. 2C). A second HLA-A2 restricted epitope, derived from the influenza type B nucleoprotein (BNP) amino acids 82 to 94 (*7*), has also been studied in a preliminary manner. The recognition of BNP82–94 by both the Y171F and Y7F mutations was abolished at concentrations up to 10^{-4} M (Fig. 2D).

Mutations at Tyr^{84} and Thr^{143} in pocket

F. Latron, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115 and Institute of Molecular Medicine, University of Oxford, United Kingdom OX3 9DU.
L. Pazmany, J. Morrison, R. Moots, A. McMichael, Institute of Molecular Medicine, University of Oxford, United Kingdom OX3 9DU.
M. A. Saper, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.
J. L. Strominger, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115 and Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.
*Current address: Biophysics Research Division, University of Michigan, Ann Arbor, MI 49109–2099.

F have also been investigated. Tyr^{84} and Thr^{143} were changed to Phe^{84} and Val^{143} (Y84F and T143V). Neither of these mutations had any detectable effect on CTL recognition of M58–66 (Fig. 3). In addition, the double mutation $Thr^{143} \rightarrow Asp^{143}$ and $Tyr^{84} \rightarrow Arg^{84}$ was constructed because these residues occur at analogous positions in a model of class II MHC glycoproteins (*18*); recognition of M58–66 or of the natural epitope formed after influenza virus infection were both abolished. Thus, the binding of peptides to class I MHC molecules and the binding of peptides to class II MHC molecules are likely to occur by distinctly different means, a conclusion also suggested by much other data.

The extra density in the cleft of HLA-A2 was presumed to be a mixture of bound peptides, because only a limited amount of detail could be seen in structural analysis at 2.6 Å (*4*), a result that was confirmed by direct analysis of this peptide mixture (*13*). The hydroxyls of Tyr^{7}, Tyr^{171}, Tyr^{84}, and Thr^{143} are in close proximity to the two ends of the electron density in the binding site of HLA-A2 (Fig. 1B) as they are in HLA-B27. The four conserved tyrosine hydroxyls in pocket A are strong hydrogen bond donors or acceptors, as are the threonines and tyrosines in pocket F, and may have a role in the binding of the amino and carboxyl ends of all peptides to MHC molecules. The distinctly different effects of the mutations in pocket A and pocket F suggest a critical function for the hydrogen bonding system in pocket A in binding the NH_3^+ terminal of all nonamers.

Loss of a single hydroxyl group at either Tyr^{171} or Tyr^{7} resulted in a two-log reduction in cytotoxic killing stimulated by either of two peptides. The loss of the Tyr^{171} hydroxyl group probably results in the elimination of one hydrogen bond to the NH_2-terminal NH_3^+ of peptide nonamers, as seen in the peptides bound to HLA-B27, or with long peptides to the amide group of the most NH_2-terminal peptide amino acid that fits into the site (for example, that between amino acids 57 and 58 in a matrix peptide). Tyr^{171} also appears to hydrogen bond to Tyr^{59}, which in turn may be accessible to solvent. Long peptides, such as M57–68, have only relatively weak affinity for HLA molecules (*16, 19*), and the mutations Y7F and Y171F do not have a further large effect on its recognition. Tyr^{7} is also hydrogen-bonded to the NH_2-terminal NH_3^+ of the peptide, as shown in crystallographic studies of H-$2K^b$ and by high resolution refinement of HLA-B27 (*9*).

Lengthening the peptide at both ends also reduced the titration value in the CTL assay by several logs (*16*), the same effect as the mutations Y171F or Y7F. These changes caused the loss of a single hydrogen bond in the A pocket that might have been expected to result in a loss in binding energy of only a few kilocalories per mole. However, the large reduction in CTL recognition indicates that these hydrogen bonds play a critical role in peptide binding, probably because the NH_2-terminal NH_3^+ is buried in the A pocket and is not accessible to water. Thus, these hydrogen bonds would make a relatively large contribution to the binding energy (*20*). The

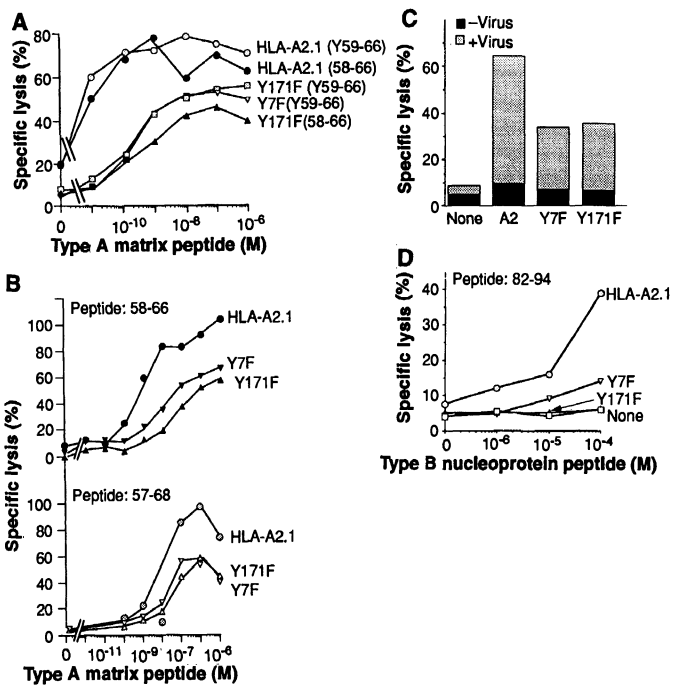

Fig. 2. Effects of mutations Y171F and Y7F in pocket A of HLA-A2 on recognition of influenza virus peptide epitopes by CTL clones. (**A**) and (**B**) Effects of the mutations Y171F and Y7F on recognition of the epitopes M58–66 and M57–68 derived from Type A influenza virus matrix protein by the HLA-A2.1 restricted JM-49 CTL line. Mutations at predefined residues in the HLA-A2.1 gene (HLA-A*0201) were obtained by site-directed mutagenesis using the procedure described (*2*). Mutated genes were subcloned into the pSVneo plasmid after linearization with Bam HI, and transfected into the HLA-A and -B negative B lymphoblastoid cell line C1R using Lipofectin. Transfected cells were selected and maintained in medium containing G418 sulfate (800 μg/ml). Cell surface expression was assessed by analytical flow cytometry with the HLA-A2 specific monoclonal antibodies BB7.2, MA2.1, 4B3, and PA2.1 by indirect immunofluorescence. Stable high expressor transfected cells of each mutated HLA-A2 gene were selected by fluorescence-activated cell sorting. Site-specific mutations at amino acid residue 7 or 171 (Y171F or Y7F) were used. Nontransfected C1R cells (labeled "None") and C1R cells transfected with the wild type HLA-A*0201 gene acted as negative and positive controls, respectively. The CTL line (*13, 16*) was grown on autologous feeder cells pulsed with M57–68 and IL-2. These experiments were done at least ten times with four different CTL lines from two individuals (J.M. and C.M.), with the same results. Cytotoxic T cell assay was quantified by a 4-hour ^{51}Cr-release assay. Target cells were labeled with sodium ^{51}Cr-chromate (Amersham, United Kingdom) for 1 hour and then mixed with the CTL at a killer/target ratio of 2/1. Matrix peptides were added to these cells immediately before the addition of CTL. The peptide epitopes used were M58–66, M57–68, and MY59–66. The substitution of Tyr for Gly at residue 58 has no effect on recognition (*16*). Peptides were synthesized by Cambridge Research and Biochemicals. Supernatants were harvested after 4 hours and specific lysis was calculated from the formula: 100 × (E-M/D-M), where E is experimental release, M is release in presence of culture medium, and D is release in presence of 5% Triton X-100. (**C**) Effects of the mutations Y171F and Y7F on recognition of Type A influenza virus infected cells by the JM CTL line. C1R or transfected C1R target cells labeled with sodium ^{51}Cr-chromate (Amersham, UK) were preincubated for 1 hour with either 100 μl of influenza A virus (hemagglutinating titer of 1/1000) in the absence of fetal calf serum (FCS) or medium alone. Target cells were then washed and resuspended in RPMI 1640 with 10% FCS for 3 hours incubation at 37°C before addition to the CTL assay. (**D**) Effects of the mutations Y171F and Y7F on recognition of the influenza B virus nucleoprotein (BNP) peptide, residues 82–94, by a HLA-A2.1 restricted CTL line, AT. Target cells as in part A labeled with sodium ^{51}Cr-chromate were first incubated with BNP peptide or medium alone for 1 hour at 37°C, and then washed before addition to the CTL assay. The CTL line AT (*7*) was used at a killer/target ratio of 2/1. CTL assays were carried out as in (A).

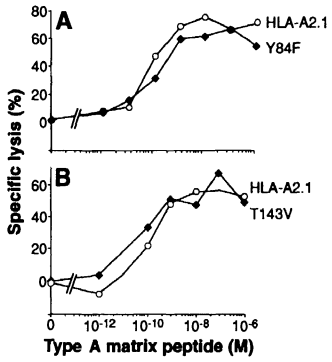

Fig. 3. Effect of the mutations Y84F and T143V in pocket F of HLA-A2 on peptide recognition. Experiments were carried out as described in Fig. 2 with Type A matrix peptide 58–66 and the CTL line JM-53.

hydrogen bonding network may result in delocalization of the positive charge on the NH_2-terminal NH_3^+, and the removal of the hydroxyl group of either Tyr^7 or Tyr^{171} could therefore destabilize the complex by disrupting this network. Thus, an intact hydrogen bonding system may be essential to bind an NH_3^+ in this region. Additionally, the binding of NH_3^+ could signal a conformational change that would lock the peptide nonamer into the site, particularly since the tyrosines involved are derived from all major structural elements forming the cleft (Tyr^{59} on the α1 helix, Tyr^{171} on the α2 helix, and Tyr^7 on the floor of the cleft). In either case, the disruption of the hydrogen bonding system would result in a large change in binding affinity (measured indirectly in this case by the CTL titration value), as observed.

In the atomic model, the hydroxyls of Tyr^{84} and Thr^{143} are positioned to hydrogen bond to the terminal carboxylate group of a short peptide, and therefore presumably to the last carbonyl group in the binding site of a longer peptide (5). The failure of the mutations Y84F and T143V to affect CTL recognition (Fig. 3) is not an unexpected result of the removal of a single hydrogen bond and may suggest a role in the mutants for water molecules in replacing the hydrogen bonds to the negatively charged peptide carboxyl group. Approximately 20 hydrogen bonds as well as van der Waals contacts involving approximately 80 to 100 atoms in 20 to 30 MHC side chains serve to bind a peptide in the cleft of a class I MHC molecule (9). Thus, the loss of a single hydrogen bond should represent a negligible loss in binding energy. Lys^{146} also appears to interact with the terminal carboxylate. A major difference between pockets A and F is the presence in the F pocket of the terminal NH_3^+ of Lys^{146} to neutralize the terminal carboxylate, whereas the A pocket does not contain a carboxylate to neutralize the terminal NH_3^+ of the nonapeptide directly (although a long range salt bridge to Glu^{63} is mediated through Tyr^7 and a water molecule) (9). This difference may also contribute to the functional importance of the hydrogen bonding network in the A pocket.

REFERENCES AND NOTES

1. P. J. Bjorkman *et al.*, *Nature* **329**, 506 (1987).
2. ———, *ibid.*, p. 512.
3. T. Garrett, M. A. Saper, P. J. Bjorkman, J. L. Strominger, D. C. Wiley, *ibid.* **342**, 692 (1989).
4. M. A. Saper, P. J. Bjorkman, D. C. Wiley, *J. Mol. Biol.* **219**, 277 (1991).
5. D. R. Madden, J. C. Gorga, J. L. Strominger, D. C. Wiley, *Nature* **353**, 321 (1991).
6. A. J. McMichael, F. M. Gotch, J. Santos-Aguado, J. L. Strominger, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9194 (1988).
7. P. A. Robbins *et al.*, *J. Immunol.* **143**, 4098 (1989).
8. J. Santos-Aguado, M. Crimmins, S. J. Mentzer, S. J. Burakoff, J. L. Strominger, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8936 (1989).
9. Crystallographic studies of H-$2K^b$ are by D. Fremont, E. A. Stura, M. Matsumara, P. A. Peterson, and I. A. Wilson, *Science* **257**, 919 (1992); and studies of HLA-B27 are by D. Madden, J. Gorga, J. L. Strominger, and D. C. Wiley (*Cell*, in press).
10. F. Latron *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11325 (1991).
11. G. M. Van Bleek and S. G. Nathenson, *Nature* **348**, 213 (1990).
12. O. Rotzschke *et al.*, *ibid.*, p. 252.
13. K. Falk, O. Rotzsche, S. Stevanovic, G. Jung, H.-G. Rammensee, *ibid.* **351**, 290 (1990).
14. F. Gotch, J. Rothbard, K. Howland, A. Townsend, A. J. McMichael, *ibid.* **326**, 653 (1987).
15. M. A. Bednarek *et al.*, *J. Immunol.* **147**, 4047 (1991).
16. J. Morrison *et al.*, *Eur. J. Immunol.* **22**, 903 (1992).
17. H. C. Bodmer, F. M. Gotch, A. J. McMichael, *Nature* **337**, 653 (1989).
18. J. Brown *et al.*, *ibid.* **332**, 845 (1988).
19. V. Cerundolo *et al.*, *Eur. J. Immunol.* **21**, 2069 (1991).
20. A. R. Fersht *et al.*, *Nature* **314**, 235 (1985).

21. We thank D. Wiley, D. Madden, and I. Wilson for many helpful discussions. Supported by grants from the National Institutes of Health, CA 47554 and AI 20182.

4 May 1992; accepted 23 July 1992

Differential Display of Eukaryotic Messenger RNA by Means of the Polymerase Chain Reaction

Peng Liang* and Arthur B. Pardee

Effective methods are needed to identify and isolate those genes that are differentially expressed in various cells or under altered conditions. This report describes a method to separate and clone individual messenger RNAs (mRNAs) by means of the polymerase chain reaction. The key element is to use a set of oligonucleotide primers, one being anchored to the polyadenylate tail of a subset of mRNAs, the other being short and arbitrary in sequence so that it anneals at different positions relative to the first primer. The mRNA subpopulations defined by these primer pairs were amplified after reverse transcription and resolved on a DNA sequencing gel. When multiple primer sets were used, reproducible patterns of amplified complementary DNA fragments were obtained that showed strong dependence on sequence specificity of either primer.

Higher organisms contain about 100,000 different genes, of which only a small fraction, perhaps 15%, are expressed in any individual cell. It is the choice of which genes are expressed that determines all life processes—development and differentiation (*1*), homeostasis, response to insults, cell cycle regulation (*2*, *3*), aging, and even programmed cell death. The course of normal development as well as the pathological changes that arise in diseases such as cancer (*4*), whether caused by a single gene mutation or a complex of multigene effects, are driven by changes in gene expression. Altered gene expression lies at the heart of the regulatory mechanisms that control cell biology. Comparisons of gene expression in different cell types provide the underlying information we need to analyze the biological processes that control our lives.

Current methods to distinguish mRNAs in comparative studies rely largely on the subtractive hybridization technique (*5*). A fingerprinting technique for mRNAs by two-dimensional (2-D) electrophoresis, such as has been used extensively in detecting cellular protein species (*6*), would be very useful. Reproducibility should be sufficient so that side-by-side comparisons of the mRNAs from different cells are possible. Furthermore, the identified spots should be usable for identifying and isolating the corresponding genes, mRNAs, or cDNAs. When protein gels were used frustration often followed because of the inability to obtain enough of the identified proteins for molecular characterization (*7*).

Our method is directed toward the identification of differentially expressed genes among the approximately 15,000 individual mRNA species in a pair of mammalian cell populations (*8*), and then recovering their cDNA and genomic clones. The general strategy is to amplify partial cDNA sequences from subsets of mRNAs by reverse transcription and the polymerase chain reaction (PCR). These short sequences are then displayed on a sequencing gel. Pairs of primers are selected so that each will amplify DNA from about 50 to 100 mRNAs because this number is optimal for display on the gel.

Selection of 3′ primers takes advantage of the polyadenylate [poly(A)] tail present on most eukaryotic mRNAs (*9*) to anchor the primer at the 3′ end of the mRNA, plus two additional 3′ bases. A primer such as 5′-$T_{11}CA$ would allow anchored annealing to mRNAs containing TG located just upstream of their poly(A) tails (*10*). By probability this primer will recognize one-twelfth of the total mRNA population because there are 12 different combinations of the last two 3′ bases, omitting T as the

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, and Division of Cell Growth and Regulation, Dana-Farber Cancer Institute, Boston, MA 02115.

*To whom correspondence should be addressed.