ENZYME CATALYSIS AND REGULATION:

Substrate Specificity of Human Collagenase 3 Assessed Using a Phage-displayed Peptide Library

Su-Jun Deng, D. Mark Bickett, Justin L. Mitchell, Millard H. Lambert, R. Kevin Blackburn, H. Luke Carter III, Jennifer Neugebauer, Gregory Pahel, Michael P. Weiner and Marcia L. Moss

J. Biol. Chem. 2000, 275:31422-31427.
doi: 10.1074/jbc.M004538200 originally published online July 21, 2000

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 40, Issue of October 6, pp. 31422–31427, 2000

Printed in U.S.A.

Substrate Specificity of Human Collagenase 3 Assessed Using a Phage-displayed Peptide Library*

Received for publication, May 25, 2000, and in revised form, July 6, 2000
Published, JBC Papers in Press, July 21, 2000, DOI 10.1074/jbc.M004538200

Su-Jun Deng‡§, D. Mark Bickett¶, Justin L. Mitchell¶, Millard H. Lambert‖, R. Kevin Blackburn‖, H. Luke Carter III‡, Jennifer Neugebauer‖, Gregory Pahel‡, Michael P. Weiner, and Marcia L. Moss¶**‡‡**

*From the Departments of ‡Molecular Sciences, ¶Molecular Biochemistry, ‖Structural Chemistry, and **Genomic Sciences, Glaxo Wellcome Research and Development, Research Triangle Park, North Carolina 27709*

The substrate specificity of human collagenase 3 (MMP-13), a member of the matrix metalloproteinase family, is investigated using a phage-displayed random hexapeptide library containing 2×10^8 independent recombinants. A total of 35 phage clones that express a peptide sequence that can be hydrolyzed by the recombinant catalytic domain of human collagenase 3 are identified. The translated DNA sequence of these clones reveals highly conserved putative P1, P2, P3 and P1′, P2′, and P3′ subsites of the peptide substrates. Kinetic analysis of synthetic peptide substrates made from human collagenase 3 selected phage clones reveals that some of the substrates are highly active and selective. The most active substrate, 2,4-dinitrophenyl-GPLGMRGL-NH_2 (CP), has a k_{cat}/K_m value of 4.22×10^6 M^{-1} s^{-1} for hydrolysis by collagenase 3. CP was synthesized as a consensus sequence deduced from the preferred subsites of the aligned 35 phage clones. Peptide substrate CP is 1300-, 11-, and 820-fold selective for human collagenase 3 over the MMPs stromelysin-1, gelatinase B, and collagenase 1, respectively. In addition, cleavage of CP is 37-fold faster than peptide NF derived from the major MMP-processing site in aggrecan. Phage display screening also selected five substrate sequences that share sequence homology with a major MMP cleavage sequence in aggrecan and seven substrate sequences that share sequence homology with the primary collagenase cleavage site of human type II collagen. In addition, putative cleavage sites similar to the consensus sequence are found in human type IV collagen. These findings support previous observations that human collagenase 3 can degrade aggrecan, type II and type IV collagens.

Human collagenase 3 (MMP-13)[1] is a member of the collagenase subfamily of matrix metalloproteinases (MMPs) (1). Several potential roles for this enzyme have been proposed, including aggrecan degradation associated with osteoarthritis (2), cleavage of type II collagen in osteoarthritic cartilage explants (3), and in tumor progression and metastasis (1). Its expression is also detected in human T lymphocytes (4). Human collagenase 3 degrades type II collagen, the primary collagen in articular cartilage, more than 10 times as efficiently as collagenase 1, another MMP known to be up-regulated in arthritic diseases (5). In addition, collagenase 3 cleaves soluble type II collagen 5 to 6 times more efficiently than type I or type III collagen (6). Furthermore a preferential collagenase 3 inhibitor significantly reduces collagen degradation from unstimulated osteoarthritic cartilage (3). These cumulative data suggest that collagenase 3 may play a significant role in the cleavage of type II collagen in arthritic cartilage.

The specificity of this enzyme for aggrecan and other collagenous substrates has also been investigated. Collagenase 3 cleaves aggrecan at the same major and minor sites identified for other members of the MMP family as well as at a novel site (7). In addition, collagenase 3 can degrade fibronectin and type IV, IX, X, and XIV collagens (8).

MMPs are considered to be potential therapeutic targets. A highly active and selective substrate is a prerequisite for screening small molecule inhibitors in high throughput screening. Knowledge of the sequence specificity can facilitate the design and optimization of potent and selective inhibitors. Synthetic peptide substrates based on physiological cleavage site sequences are routinely used for enzymatic assays. These substrates can often be improved because physiologically based peptide substrates may not be the optimal substrate (9, 10). To optimize protease substrates, a number of chemical methods have been developed. These include using defined substrate mixtures (11, 12) and immobilized peptide libraries (13). Recently two biological systems using phage display have been developed to quickly search and decode a large peptide library for substrate optimization (14, 15). The first system uses monovalent substrate phage display to study substrate specificity of proteases (14). The second system uses polyvalent substrate phage display to more quickly optimize protease substrates (15, 16). These chemical and biological methods may also be useful in searching for substrates of proteases where the natural physiological substrate is unknown.

In this study, we have screened a phage-displayed random hexapeptide library of 2×10^8 members using the catalytic domain of recombinant human collagenase 3. We have identified 35 phage clones that express a hexapeptide sequence cleaved by the enzyme. We searched sequence data bases for potential collagenase 3 cleavage sites by running FASTA and BLAST searches with the hexapeptide phage sequences as queries. These searches identified potential collagenase 3

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Dept. of Molecular Sciences, Glaxo Wellcome Inc., P. O. Box 13398, 5 Moore Dr., Research Triangle Park, NC 27709. Tel.: 919-483-3819; Fax: 919-483-0368; E-mail: sjd20506@glaxowellcome.com.

‡‡ Present address: Dept. of Biochemistry and Biophysics, Lineberger Comprehensive Cancer Center 227, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

[1] The abbreviations used are: MMP, matrix metalloprotease; MMP-13, human collagenase 3; HPLC, high performance liquid chromatography; mAb, monoclonal antibody; TGF, transforming growth factor; DTT, dithiothreitol; CP, 2,4-dinitrophenyl-GPLGMRGL-NH_2.

cleavage sites in type IV collagen and the latency-associated peptide of TGF-β−3. Using collagenase 3 and other closely related MMPs, we have also carried out kinetic studies for synthetic peptide substrates synthesized based on the hexapeptide sequence from a selected number of the 35 phage clones. These synthetic peptides are shown to be highly sensitive and selective to human collagenase 3.

EXPERIMENTAL PROCEDURES

Reagents—Escherichia coli K91 was obtained from Marc Navre of Affymax. *E. coli* TG1 was purchased from Amersham Pharmacia Biotech. All restriction enzymes were purchased from New England Biolabs (Beverly, MA). ECL system was purchased from Amersham Pharmacia Biotech. Pansorbin cells were purchased from Calbiochem. Monoclonal antibody 3-E7 was purchased from Gramsch Laboratories (Schwabhausen, Germany). Monoclonal antibody mAb 179 was prepared using standard methods. Dimethyl sulfoxide, *p*-aminophenylmercuric acetate, HEPES, and Brij 35 were purchased from Sigma.

Purification of Human Recombinant Collagenase 3—The catalytic domain of human collagenase 3 was prepared by PCR cloning from the original pNOT3α clone (1) using the 5-prime oligonucleotide 5′-CCCGGGCATATGTACAATGTTTTCCCTCGTACGCTTAAATGGTCCAAAACGAATTTAACCTACAGA and the 3-prime oligonucleotide 5′-CCCGGGATCCTTAACCATAGAGAGACTGAATCCCTTGTACATCGTCATCAGGAAG. Base pairs 34–36 of the 5-prime oligo change an ATG to CTT, which removes a problematic second start site, whereas base pair 29 of the 3-prime oligo changes a G to an A, which removes a naturally occurring *Bam*HI site. The final product was cloned via *Nde*I and *Bam*HI into pET11b. A 10-liter broth culture of *E. coli* BL21[DE3] bearing pETcoll3 was grown at 37 °C to an A_{600} of 1, shifted to 16 °C, and induced with isopropyl-1-thio-β-D-galactopyranoside, and growth was continued for 4 h until harvest. The *E. coli* cell pellet was suspended in 2 ml of lysis buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 10 mM EDTA, 1 mM DTT, 50 μg/ml APMSF (4-amidinophenyl)methanesulfonylfluoride), 20 μg/ml aprotinin, 10 μg/ml bestatin, 5 μg/ml pepstatin A, and 10 μg/ml leupeptin)/g of wet weight cells. Cells were lysed with two passes through a French press. Inclusion bodies were spun at 48,000 × *g* and washed twice with 1.5 ml of Triton buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 10 mM EDTA, 1 mM DTT, and 0.5% Triton X-100)/g of cells and once with 1.5 ml of low salt buffer (50 mM Tris, pH 7.5, 10 mM EDTA, 1 mM DTT, and 0.5% Triton X-100)/g of cells. The human recombinant collagenase 3 was solubilized with 1 ml of extraction buffer (50 mM Tris, pH 7.5, 1 mM DTT, and 6 M urea)/g of cells. Protein was applied to an anion exchange column (PerSeptive Biosystems Poros HQ20) equilibrated with column buffer (20 mM Tris, pH 7.5, 1 mM DTT, and 6 M urea) and eluted with a linear NaCl gradient (0 to 300 mM). Fractions containing collagenase 3 were identified by SDS-PAGE and pooled. Collagenase 3 was refolded at 4 °C by rapid dilution to 100 μg/ml (~45-fold dilution) in folding buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 50 μM $ZnCl_2$, and 1 mM $CaCl_2$). Unfolded or incorrectly folded protein was removed by filtration (0.2 μm). The active enzyme was concentrated and stored at −80 °C.

Construction of Control Substrate Phage—Negative collagenase 3-resistant (fTCol-3R) and positive collagenase 3-sensitive (fTCol-3S) control substrate phage clones were constructed by ligating oligonucleotides with *Kpn*I/*Xho*I-digested fAFF-TC-LIB vector (15). Oligonucleotides 5′-TCGAGCGGTGGTAGTGGTAATCCTGTTGAGCCTGCTCTGGTAC and 5′-CAGAGCAGGCTCAACAGGATTACCACTACCACCGC were used for the construction of fTCcol-3R phage. Oligonucleotides 5′-TCGAGCGGTGGTAGTGGTCCGCTGGGTCTGTGGGCCCGTCTGGTAC and 5′-CAGACGGGCCCACAGACCCAGCGGACCACTACCACCGC were used for the construction of fTCol-3S phage. The displayed peptide sequence of fTCol-3R was described previously (15). The substrate sequence of fTCol-3S was based on peptide substrate 2,4-dinitrophenyl-PLGLWAR for MMPs (17, 18).

Phage-displayed Hexapeptide Library Screening—The construction of phage-displayed hexapeptide library fTC-LIB-N6 with 2×10^8 individual sequence members was as described previously (15). A total of 2×10^{10} fTC-LIB-N6 phage were digested at 37 °C for 1 h by 2 nM collagenase 3 in a 250-μl reaction mixture containing 1× MMP buffer (50 mM Tris-HCl, 5 mM $CaCl_2$, 10 μM $ZnSO_4$, 200 mM NaCl, pH7.6) and 0.005% Brij 35. The reaction was terminated by the addition of 0.5 M EDTA to a final concentration of 5 mM. Then 100 μg of mAb 179 and 10 μg of mAb 3-E7 was added to the terminated reaction mixture followed by the addition of bovine serum albumin to 0.1%. Both of the monoclonal antibodies were used to capture uncleaved phage as described previously (15). After 30 min of incubation on ice, 100 μl of Pansorbin cells were added, and the reaction mixture was rotated at 4 °C for 1 h. The mixture was centrifuged for 2 min at 12,000 × *g*, and the supernatant was recovered. The adsorption of uncleaved phage-mAb complex to Pansorbin cells was repeated once. Phage in the final supernatant were transfected into *E. coli* K91 and amplified overnight at 37 °C. Phage in a small aliquot of the final supernatant was titered on *E. coli* K91. The above screening was performed for five rounds. Single phage clones from the titer plates were picked randomly and grown in 4 ml of LB for isolation of single-stranded DNA and for Western blotting analysis of phage.

Construction of Site-directed Mutant Phage, Consensus Phage, and Aggrecan-derived Phage—C2–12P3A, NF, DL, PV and consensus CP phage clones were constructed by inserting oligo cassettes into *Kpn*I/*Xho*I-digested fAFF-TC-LIB vector and confirmed by DNA sequencing. NF phage was made based on the major MMP site (PEN-FFG), and DL phage was made based on the minor MMP site (SED-LVV) in aggrecan. PV phage was constructed on the basis of a novel human collagenase 3 cleavage site (VKP-VFE), which was not observed for other MMPs in aggrecan (7). Oligonucleotides 5′-TCGAGCGGTGGTAGTGGTGGTGCTTTGGGGTTGTCGCTGGTAC and 5′-CAGCGACAACCCCAAAGCACCACCACTACCACCGC were used for the construction of C2–12P3A. Oligonucleotides 5′-TCGAGCGGTGGTAGTGGTCCGGAAAACTTCTTCGGTCTGGTAC and 5′-CAGACCGAAGAAGTTTTCCGGACCACTACCACCGC were used for the construction of NF phage clone. Oligonucleotides 5′-TCGAGCGGTGGTAGTGGTTCTGAAGACCTGGTTGTTCTGGTAC and 5′-CAGAACAACCAGGTCTTCAGAACCACTACCACCGC were used for the construction of DL phage clone. PV phage was made by 5′-TCGAGCGGTGGTAGTGGTGTTAAACCGGTTTTCGAACTGGTAC and 5′-CAGTTCGAAAACCGGTTTAACACCACTACCACCGC. Consensus CP phage clone was made with oligonucleotides 5′-TCGAGCGGTGGTAGTGGTCCTTTGGGTATGCGTGGTCTGGTAC and 5′-CAGACCACGCATACCCAAAGGACCACTACCACCGC.

Western Blotting Analysis and DNA Sequencing—Phage were precipitated by adding 200 μl of 20% polyethylene glycol, 2.5 M NaCl to 1 ml of phage supernatant and centrifuging at 12,000 × *g* for 15 min. Phage pellets were dissolved in 100 μl of Tris-buffered saline. Three different enzymatic reactions were performed. For the first reaction, phage was incubated in 50 μl of 1× MMP buffer and 0.005% Brij 35 at 37 °C for 1 h, but enzyme was not added. For the second reaction, phage was treated as described for the first reaction except 0.4 nM collagenase 3 was added. For the third reaction, phage were treated the same way as described for the first reaction except 0.4 nM collagenase 3 and 0.05 μl of 10 μM GW9471 inhibitor were combined (19). EDTA was added to 5 mM to terminate enzymatic reactions. Each reaction mixture of 10 μl was analyzed on 12% SDS-polyacrylamide gels. Substrate peptide-phage gene gIII fusions were detected by mAb 179. Phage cleaved by human collagenase 3 but whose cleavage was inhibited by the broad spectrum MMP inhibitor GW9471 was chosen for further study. Single-stranded DNA from individual phage clones was prepared using QIAprep M13 BioRobot kit in a Bio Robot 9600 system (Qiagen, CA). The phage substrate DNA sequences were determined by an ABI automated fluorescent DNA sequencer.

Collagenase 1 (MMP-1), Stromelysin 1 (MMP-3), and Gelatinase B (MMP-9) for Selectivity Studies—A 19-kDa collagenase 1 catalytic domain was expressed in *E. coli.* Purification was described previously (20). Full-length stromelysin 1 was also expressed in *E. coli* and purified from inclusion bodies using a protocol similar to that used for collagenase 1. Truncated progelatinase B (lacking the C-terminal domain) was expressed from baculovirus infected *Trichoplusia ni* cells and purified from the media using gelatin-agarose affinity chromatography. Purity was greater than 90%, as assessed by SDS gel electrophoresis.

Estimation of Specificity Constants k_{cat}/K_m—All synthetic 2,4-dinitrophenyl peptide substrates were ordered from Synpep (Dublin, CA) and prepared as 10 mM stock solutions in dimethyl sulfoxide. A substrate concentration of 10 μM was used for all enzymatic reactions. Before the assay, collagenase 3 was diluted 100-fold to a 250 nM stock in 200 mM NaCl, 50 mM Tris, 5 mM $CaCl_2$, 10 μM $ZnSO_4$, 0.01% Brij 35, pH 7.5 (assay buffer). Hydrolysis reaction concentrations were adjusted from 0.1 to 10 nM by further dilution. The samples were incubated at 37 °C for 10 to 180 min. The reactions were terminated with the addition of an equal volume of 0.1% heptafluorobutyric acid. Peptide substrate and product were separated by reverse phase HPLC, and the peaks were evaluated by monitoring absorbance at 350 nm. Cleavage sites in synthetic peptides were determined by liquid chromatography-mass spectrometry analysis of each reaction mixture.

The specificity constants k_{cat}/K_m were also determined for collagenase 1, stromelysin 1, and gelatinase B. Collagenase 1 was refolded by dilution to 5 μM with extensive dialysis against 50 mM Tris/HCl, pH

TABLE I

Translated amino acid sequences of human collagenase 3-selected phage clones from a random hexapeptide substrate phage library

+++, strong cleavage; ++, some cleavage; +, less cleavage; ND, no cleavage of phage by human collagenase 3 detected by Western blotting as described under "Experimental Procedures." Phage clones were named in such a way that C2–12 represented clone 12 from round 2 screening. Amino acids in bold are preferred residues and are represented in the consensus sequence CP.

Phage clone	Amino acid sequence								Cleavage
			P3	P2	P1	P1′	P2′	P3′	
fTCol-3R		N	**P**	V	E	P	A		ND
fTCol-3S			**P**	**L**	**G**	L	W	A R	+
C2–12		G	**P**	**L**	**G**	L	S		+++
C3–9		G	**P**	**L**	**G**	L	**K**		+++
C2–6	V	R	**P**	**L**	**G**	I			++
C5–19/5–24		A	**P**	**L**	**G**	**M**	S		+++
C3–10			**P**	**L**	S	I	S	**G**	+++
C3–15			**P**	**L**	Q	F	**R**	**G**	+++
C3–19			**P**	**L**	T	**M**	M	**G**	+++
C5–10			**P**	**L**	P	**M**	**R**	M	++
C5–9		A	**P**	**L**	A	**M**	**R**		+++
C5–12			**P**	**L**	S	I	Q	D	+++
C3–20			**P**	**L**	S	F	Q	**G**	+++
C3–26	V	L	**P**	**L**	P	**M**			+++
C4–21			**P**	**L**	S	L	L	S	+++
C3–18			**P**	A	**G**	L	S	D	+++
C4–8			**P**	R	**G**	L	V	A	+++
C3–12			**P**	K	**G**	L	**R**	A	+++
C4–18			**P**	Y	**G**	**M**	**R**	A	+++
C4–13/4–27			**P**	V	A	L	**K**	A	+++
C3–16/4–1			**P**	K	**G**	V	Y	S	+++
C4–19			**P**	K	**G**	I	T	S	+++
C4–24			**P**	S	**G**	I	**H**	V	+++
C5–27			**P**	F	**G**	F	**K**	S	+++
C4–4	G	P	**P**	M	S	L			+++
C3–33			**P**	I	N	L	**H**	**G**	+++
C2–16			**P**	H	P	L	F	L	++
C2–2			**P**	S	E	L	**K**	**G**	++
C2–7	M	A	**P**	Y	A	L			+++
C4–14/5–25			**P**	F	A	F	Q	**G**	+++
C2–14			**P**	S	A	Y	**H**	S	+++
C5–15			**P**	M	S	Y	N	**A**	+++
C2–27			**P**	A	E	I	V	**A**	+++
C4–1			**P**	M	E	**M**	V	E	++
C2–9	V	T	**P**	Y	N	**M**			+++
C2–23		P	**P**	R	A	I	R		+++
C2–22			**P**	R	P	F	N	Y	+++
Consensus CP			**P**	**L**	**G**	**M**	**R**	**G**	+++

TABLE II

Cleavage of control phage, site-directed mutant phage, and aggrecan-derived phage

+++, strong cleavage; ++, some cleavage; +, less cleavage; ND, no cleavage of phage by human collagenase 3 detected by Western blotting as described under "Experimental Procedures." DL and NF phage were constructed based on minor and major MMP cleavage sites in aggrecan. PV phage was constructed based on a novel site not observed for other MMPs in aggrecan. DL, NF, and PV phage were named according to their P1 and P1′ residues (7).

Phage clone		Subsite							Cleavage
		P3	P2	P1	P1′	P2′	P3′		
fTCol-3R	N	P	V	E	P	A			ND
fTCol-3S		P	L	G	L	W	A	R	+
C2–12	G	P	L	G	L	S			+++
C2–12P3A	G	A	L	G	L	S			ND
DL		S	E	D	L	V	V		ND
NF		P	E	N	F	F	G		++
PV		V	K	P	V	F	E		ND

7.5, containing 200 mM NaCl, 1 mM $CaCl_2$, and 50 μM $ZnCl_2$ (refolding buffer). Hydrolysis reactions were performed using 1–10 nM enzyme in assay buffer. Stromelysin 1 was refolded by diluting to 15 nM in assay buffer. After 1 h at room temperature, the enzyme was concentrated to a 150 nM stock solution using an Amicon CentriPrep 30 concentrator. Stromelysin 1 assay concentrations ranged from 5 to 50 nM. Gelatinase B was diluted to a 2.5 μM stock in assay buffer and activated for 12 h

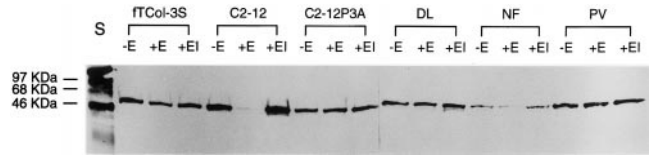

FIG. 1. **Western blot analysis of phage clones treated with recombinant human collagenase 3.** Approximately 10^8 phage colony-forming units from collagenase 3-positive control phage fTCol-3S and phage clones C2–12, C2–12P3A, DL, NF, and PV were treated without enzyme (−*E*), with 0.4 nM recombinant human collagenase 3 (+*E*), or with 0.4 nM recombinant human collagenase 3 and 10 μM GW9471 inhibitor (+*EI*) as described under "Experimental Procedures." *S* represents ECL molecular mass standard.

TABLE III

Comparison of collagenase 3-selected substrate phage with a phage clone constructed based on major MMP cleavage site NF in aggrecan

+++, strong cleavage; ++, some cleavage of phage by human collagenase 3 detected by Western blotting as described under "Experimental Procedures." Amino acids shown in bold were observed in the major MMP cleavage site NF in aggrecan.

Phage clone	Subsite						Cleavage
	P3	P2	P1	P1′	P2′	P3′	
NF	**P**	E	N	**F**	F	**G**	++
C3–15	**P**	L	Q	**F**	R	**G**	+++
C3–20	**P**	L	S	**F**	Q	**G**	+++
C4–14/5–25	**P**	F	A	**F**	Q	**G**	+++
C5–27	**P**	F	G	**F**	K	S	+++
C2–22	**P**	R	P	**F**	N	Y	+++

with 1 mM *p*-aminophenylmercuric acetate at 37 °C. The enzyme was desalted with Bio-Rad Bio-Spin 6 columns into assay buffer. The enzyme was diluted to a final assay concentration of 1–10 nM. All enzyme concentrations were verified by active site titration with the broad spectrum MMP inhibitor, GW9471 (19). The IC_{50} for collagenase 3, collagenase 1, gelatinase B, and stromelysin 1 is 0.5, 1.2, 1.0, and 2.4 nM, respectively, for this inhibitor. A substrate concentration of 10 μM was used for each enzymatic reaction. Hydrolysis reactions and HPLC analysis were performed under the same conditions as for collagenase 3. The standard error for specificity constant k_{cat}/K_m is less than or equal to 30% based on 3 independent experiments. The specificity constant k_{cat}/K_m for 2,4-dinitrophenyl-GPLGMRGL-NH_2 (CP) was also estimated by performing enzymatic reactions at a substrate concentration of 2.5 μM, and the k_{cat}/K_m value was similar to that estimated at a substrate concentration of 10 μM.

RESULTS

A total of 2×10^{10} fTC-LIB-N6 phage from the phage-displayed random hexapeptide library was treated with the catalytic domain of human collagenase 3. Undigested phage was removed by mAb 3E-7 and mAb 179 attached to Pansorbin cells. Cleaved phage was amplified in *E. coli* K91 cells and then subjected to four more rounds of enzyme digestion and screening. This screening yielded 35 phage clones, which were shown to be more sensitive by Western blot analysis to human collagenase 3 treatment than the positive control phage fTCol-3S (Table I). DNA sequencing of the 35 phage clones and sequence alignment analysis reveal that a proline appears in all 35 clones in a putative P3 subsite, a leucine appears in 13 clones at the P2 subsite, a glycine occurs in 12 clones at the P1 subsite, a methionine occurs in 8 clones at the P1′ subsite, a basic residue arginine, lysine, or histidine occurs in 12 clones at the P2′ subsite, and a glycine occurs in 7 clones at the P3′ subsite (Table I). A consensus sequence, PLGMRG, can be deduced based on the preferred residue in each subsite position. A site-directed mutant phage C2–12P3A was constructed in which proline at the putative P3 subsite in phage clone C2–12 was replaced by alanine (Table II). Western blotting did not detect cleavage of the site-directed mutant by collagenase 3, whereas the parent phage was cleaved (Table II and Fig. 1). The k_{cat}/K_m for the synthetic peptide, however, was reduced

TABLE IV

Comparison of human collagenase 3-selected peptide substrate sequences and the primary collagenase cleavage site in human type II collagen

+++, strong cleavage; ++, some cleavage detected of phage by human collagenase 3. Amino acids shown in bold were observed in type II collagen.

Phage clone or human collagens	Subsite								Cleavage
	P4	P3	P2	P1	P1′	P2′	P3′	P4′	
Cleavage site of type IIα1 collagen	**G**	**P**	Q	**G**	**L**	A	**G**	Q	
C2–12	**G**	**P**	L	**G**	**L**	S			+++
C3–9	**G**	**P**	L	**G**	**L**	K			+++
C3–18		**P**	A	**G**	**L**	S	D		+++
C4–8		**P**	R	**G**	**L**	V	A		+++
C3–12		**P**	K	**G**	**L**	R	A		+++
C3–33		**P**	I	N	**L**	H	**G**		+++
C2–2		**P**	S	E	**L**	K	**G**		++

TABLE V

Comparison of specificity constant (k_{cat}/K_m) values for hydrolysis by MMPs of peptides synthesized based on selected substrate phage clones, a site-directed phage mutant, and aggrecan sequence

Peptide CP was made based on the consensus sequence of human collagenase 3-selected phage clones. Peptides DL, NF, and PV were made based on MMP cleavage sites and a collagenase 3 site in aggrecan. All peptides contained an N-terminal 2,4-dinitrophenyl peptide and a C-terminal amide. NT, not tested. ND, no cleavage detected. The arrow indicates the cleavage site.

Synthetic peptide sequence subsites							Derived from phage clone	k_{cat}/K_m			
P3	P2	P1	P1′	P2′	P3′			Collagenase 3	Collagenase 1	Gelatinase B	Stromelysin 1
		↓						$M^{-1}\ s^{-1}$			
G	P	L	G	M	R	G	L	Consensus CP	4.22E+06	5.14E+03	3.90E+05	3.14E+03
G	P	L	G	L	W	A	R	fTCol-3S	1.41E+06	3.06E+04	1.03E+05	4.00E+03
G	P	R	P	F	N	Y	L	C2–22	1.08E+06	5.87E+03	5.15E+04	1.37E+05
G	P	F	G	F	K	S	L	C5–27	5.11E+05	1.74E+02	1.07E+05	2.08E+03
G	P	L	G	L	S	L		C2–12	2.41E+05	3.33E+03	6.67E+04	1.17E+04
G	A	L	G	L	S	L		C2–12P3A	3.53E+04	4.25E+03	7.59E+03	2.48E+03
G	P	M	S	Y	N	A	L	C5–15	2.25E+05	5.56E+03	2.50E+04	2.28E+03
G	P	K	G	V	Y	S	L	C3–16	1.60E+05	2.92E+01	7.21E+04	4.51E+01
G	P	A	G	L	S	D	L	C3–18	1.53E+05	5.28E+03	1.50E+04	7.32E+02
G	P	E	N	F	F	G	L	NF	1.14E+05	4.44E+03	8.00E+04	2.35E+03
G	P	F	A	F	Q	G	L	C4–14	1.96E+05	4.72E+03	2.33E+05	3.27E+03
G	V	K	P	V	F	E	L	PV	ND	NT	NT	NT
G	S	E	D	L	V	V	L	DL	ND	NT	NT	NT

TABLE VI

Selectivity of collagenase 3 substrates with respect to collagenase 1, gelatinase B and stromelysin 1

Selectivity of synthetic peptide substrates for human collagenase 3 over collagenase 1, gelatinase B, and stromelysin 1 was based on the ratio of collagenase 3 k_{cat}/K_m over the k_{cat}/K_m of other MMPs.

Derived from phage clone	Selectivity		
	Collagenase 3/ Collagenase 1	Collagenase 3/ Gelatinase B	Collagenase 3/ Stromelysin 1
Consensus, CP	820	11	1300
fTCol-3S	46	14	350
C2–22	180	21	7.9
C5–27	2900	4.8	250
C2–12	72	3.6	21
C2–12P3A	8.3	4.6	14
C5–15	41	9.0	99
C3–16	5500	2.2	3600
C3–18	29	10	210
NF	26	1.4	49
C4–14	41	0.8	60

7-fold by the alanine substitution (Table V).

Phage display screening resulted in five phage clones that shared sequence homology with a major MMP cleavage site NF for aggrecan (7). All 5 clones (C3–15, C3–20, C4–14/5–25, C2–22, and C5–27) have a proline at the P3 subsite and a phenylalanine in the P1′ subsite. Three clones (C3–15, C4–14/5–25, and C3–20) have a glycine at P3′ subsite (Table III). Three phage clones, DL, NF, PV, corresponding to minor, major MMP cleavage sites, and a novel collagenase 3 site, reported for aggrecan, were constructed (Table II). Of these, only the NF phage that was based on the major MMP cleavage site for aggrecan was obviously sensitive to human collagenase 3 treatment as shown by Western blot analysis (Table II and Fig. 1).

Phage display screening selected seven phage clones (C2–12, C3–9, C3–18, C4–8, C3–12, C3–33, and C2–2) that shared sequence homology to the primary collagenase cleavage site of human type II collagen (3). All seven clones have a proline at the P3 subsite and a leucine at the P1′ subsite, five clones (C2–12, C3–9, C3–18, C4–8, and C3–12) have a glycine at the P1 subsite, and another two clones (C3–33 and C2–2) have a glycine at the P3′subsite (Table IV).

Synthetic peptide substrate CP, corresponding to the consensus sequence, showed the highest k_{cat}/K_m value of 4.22×10^6 $M^{-1}\ s^{-1}$ for hydrolysis by human collagenase 3. The k_{cat}/K_m value of peptide CP for human collagenase 3 is 1344-, 11-, and 820-fold higher than that for the MMPs stromelysin 1, gelatinase B, and collagenase 1, respectively. Peptide substrate CP has a 3-fold-improved k_{cat}/K_m value for hydrolysis by collagenase 3 when compared with the synthetic peptide substrate corresponding to positive control phage, fTCol-3S. Furthermore, the peptide substrate CP has a higher k_{cat}/K_m than described for previous substrates such as McaPChaGNva-HADpa-NH_2, McaPLGLDpaAR-NH_2 (6), and the triple helical collagen substrate, α1(I)-772-786 triple helical peptide (22). With respect to selectivity for collagenase 3 over collagenase 1 and stromelysin 1, peptide CP also has a greater relative selectivity than the peptide corresponding to control phage, fTCol-3S (Tables V and VI). The NF synthetic peptide corresponding to the major MMP cleavage site in aggrecan has a 37-fold lower k_{cat}/K_m value for hydrolysis by collagenase 3 compared with peptide CP. It has a 49-, 1-, and 26-fold human collagenase 3 selectivity over related MMPs stromelysin 1,

TABLE VII

Interesting hits from FASTA and/or BLAST search of the SwissProt and translated EMBL protein data bases for potential human collagenase 3 cleavage sites using hexapeptide phage sequences as queries

aa is an abbreviation for amino acid.

Phage clone	Query	Hits	Human protein	Putative Site	Cleavage Domain
C3–9	GPLGLK	GPMGLK	α2 Chain of Type IV collagen	aa 1000	Triple-helical region
		GPVGMK	α2 Chain of Type IV collagen	aa 881	Triple-helical region
C2–12	GPLGLS	GPIGLS	α4 Chain of Type IV collagen	aa 86	Triple-helical region
C3–16	PKGVYS	PKGVFS	Biglycan	aa 177	Mature protein
C4–19	PKGITS	PKGITS	TGF-β3	aa 150	Latency-associated peptide

gelatinase B, and collagenase 1. Substrates GPKGVYSL and GPFGFKSL, corresponding to phage clones C3–16 and C5–27, are good substrates for collagenase 3 and gelatinase B, but relatively poorer substrates for collagenase 1 and stromelysin 1. Highly selective substrates have been described (23, 24) and shown to have utility for identifying specific MMPs in biological samples (25). No hydrolysis was observed by collagenase 3 for synthetic peptides PV and DL that were synthesized based on a reported minor MMP site and a novel collagenase 3 site, respectively, in aggrecan.

We searched the SwissProt and translated EMBL protein sequence data bases for potential collagenase 3 cleavage sites by running FASTA and/or BLAST searches using the consensus sequence and each of the 35 hexapeptide phage sequences as queries. Considering the 17,000 human sequences available in March 2000, the 36 searches identified 12 different 6-residue sites that were identical to one or more of the phage sequences and numerous 6-residue sites that were similar to one or more of the phage sequences.

Eight of the identical matches occur in sequences that are poorly characterized or in hypothetical translations of DNA sequences; these are difficult to interpret. Three of the identical matches occur in proteins that are expected to be intracellular and, thus, unavailable for cleavage by collagenase 3. The one remaining identical match occurs in the sequence of TGF-β3, a well characterized extracellular protein. This match involves the sequence PKGITS in the prodomain of TGF-β3, which is identical to the sequence of phage clone C4–19 (Table VII). This match is particularly interesting since activation of TGF-β involves proteolysis by metalloproteases (26).

Several of the near matches occur in collagens and collagen-like sequences, some of which may be substrates for collagenase 3. For example, Knauper *et al.* report that collagenase 3 degrades type IV collagen, although the actual cleavage site(s) is unknown (8). The FASTA searches identified several potential cleavage sites in the α2(IV) chain of type-IV collagen, including GPMGLK (residues 997–1002, similar to phage clone C3–9), GPLGLP (residues 851–856, similar to phage clone C2–12), and GPVGMK (residues 879–884, similar to the consensus sequence). Cleavage of collagen by collagenase 1 and 2 is thought to involve local unwinding of the triple helix, with a tightly triple-helical region upstream from the cleavage site and a loosely triple-helical region downstream from the cleavage site (27). Thus, the local conformation and more distant flanking residues would presumably determine whether the α2(IV) chain is actually cleaved by collagenase 3 at any of these sites.

Additional near matches occur in proteins of the extracellular matrix. For example, biglycan, a small proteoglycan similar to decorin and found in connective tissue (28), contains a PKGVFS sequence (residues 175–180) that is almost identical to phage clone C3–16 (PKGVYS). The corresponding synthetic peptide was hydrolyzed by both collagenase 3 and gelatinase B (Table VI), suggesting that biglycan might be a substrate for either or both of these metalloproteases.

DISCUSSION

It has been proposed that collagenase 3 is associated with osteoarthritis, rheumatoid arthritis, and tumor progression and metastasis (2, 3, 5). To gain further insights into the substrate specificity of this enzyme, we screened a phage-displayed random hexapeptide library. We have identified 35 phage clones that are cleaved by the recombinant catalytic domain of collagenase 3. The presumed cleavage site and corresponding P3, P2, P1, P1′, P2′, and P3′ positions were identified from a sequence alignment. All of the selected peptide substrates have proline at the P3 position and lipophilic amino acids at P1′. Most of the selected substrates have lipophilic amino acids at P2 and small amino acids at P1 and P3′. Many of the substrates have a basic amino acid at P2′. Our observation that the site-directed mutant C2–12P3A is 7-fold less sensitive to collagenase 3 treatment indicates the important role of the P3 proline in substrate specificity (Fig. 1, Tables II and V). Earlier studies (24) showed that interstitial collagenase (MMP-1) and neutrophil collagenase (MMP-8) also have a preference for proline at P3 and for lipophilic residues at P1′. However, the preference for proline at P3 appears to be stronger in collagenase 3 than in the other two collagenases. Collagenase 3 and neutrophil collagenase can accommodate both aromatic and aliphatic amino acids at P1′, whereas interstitial collagenase has a stronger preference for aliphatic amino acids (24). This is consistent with x-ray structures, which show that neutrophil collagenase and collagenase 3 have deep lipophilic S1′ pockets, whereas the S1′ pocket in interstitial collagenase is partially blocked by an arginine side chain (29).

Screening selected seven substrate sequences with sequence homology to the primary collagenase cleavage site in type II collagen (PQG-LAG). This is in agreement with previous reports that human collagenase 3 cleaves type II collagen (3). Previous work (7) showed that collagenase 3 cleaves the interglobular domain of aggrecan at a major MMP site (PEN-FFG), a minor MMP site (SED-LVV), and a novel collagenase 3 site (VKP-VFE). Our phage display screening selected five substrates similar to the PEN-FFG site (Table III) but none similar to the SED-LVV or VKP-VFE sites. Western blot analysis showed that substrate phage clones DL and PV, corresponding to the minor MMP cleavage site and the novel collagenase 3 cleavage site, were not sensitive to collagenase 3 treatment (Fig. 1). No turnover is observed for the corresponding synthetic peptides DL and PV (Table V). Cleavage of these sites by collagenase 3 evidently requires additional flanking residues and/or the specific conformation adopted in the interglobular domain of aggrecan.

Most of the synthetic peptides corresponding to phage clones have reasonable specificity constants. The synthetic peptide substrate corresponding to the consensus sequence CP has the highest collagenase 3 k_{cat}/K_m value, 4.22×10^6 M^{-1} s^{-1}. This k_{cat}/K_m is 1300-, 11-, and 820-fold better than that of related MMPs stromelysin 1, gelatinase B, and collagenase 1 (Tables V and VI). Interestingly, the fTCol-3S peptide substrate GPLGLWAR has a k_{cat}/K_m value of 1.4×10^6 M^{-1} s^{-1} for human

collagenase 3. However, the corresponding control phage is only weakly hydrolyzed as assessed by Western blot. It is possible that steric interference of phage protein on the displayed control substrate sequence may have caused the reduction of turnover. Therefore it is possible that some of the substrate sequences in the phage-displayed hexapeptide library may not have been selected because of steric interference and/or the lesser sensitivity of Western blotting. The k_{cat}/K_m value for collagenase 1 is also significantly higher for GPLGLWAR than what has been reported previously in the literature (17, 18). The explanation is unclear but may be due to differences in detergent concentration or buffer composition as well as potential differences related to the use of a truncated construct.

This work demonstrates the utility of phage display in identifying highly active and selective substrates. The most active substrate, CP, and selective substrates, C3–16 and C5–27, may be useful in the screening and design of selective collagenase 3 inhibitors for cancer and arthritic diseases. In addition, the more selective substrates may be used to monitor activity of collagenase 3 in disease samples. An example of the application of a selective substrate was reported for stromelysin, where levels of the enzyme were determined in rheumatoid synovial fluid by a fluorogenic assay using a selective peptide substrate (25). Phage display can also help define the substrate specificity of a protease more completely. In general, this could be useful in recognizing physiologically relevant substrates.

In this particular case, data base searches identified sites in many proteins with similarity to one or more of the phage query sequences. However only a few proteins showed identical or one conservative mismatch sequences to the phage query sequences. Most of these sites occur in poorly characterized or hypothetical proteins or in proteins that are presumably intracellular and, thus, unavailable to collagenase 3. However, several sites seem plausible enough to suggest further experiments. Of these, the most interesting is TGF-β3, which contains a PKGITS sequence identical to that of clone C4–19. TGF-β3 is a potent angiogenesis factor that also plays a role in wound healing and bone homeostasis (30). TGF-β is released from cells as an inactive complex consisting of the mature TGF-β protein bound to its cleaved prodomain, which is then referred to as the TGF-β latency-associated protein. (31). This inactive form of TGF-β can be activated by enzymatic or chemical treatment, allowing it to bind to its receptor (32). Recently, Yu and Stamenkovic (26) discovered that cell surface gelatinase B proteolytically activated TGF-β1, 2, and 3, thereby acting as an angiogenesis factor and promoting tumor invasion and metastasis. The PKGITS sequence identified by our phage display work lies near the middle of the TGF-β3 prodomain. Collagenase 3 might thus cleave the TGF-β3 latency-associated protein, possibly breaking up the complex and thereby activating TGF-β3. The most similar synthetic peptide contains the PKGVYS sequence, corresponding to phage clone C3–16. This synthetic peptide was cleaved efficiently by collagenase 3 and gelatinase B, suggesting that either of these enzymes might cleave TGF-β3 at this position. Of course, this site might be inaccessible in the three-dimensional structure of the complex. Note also that activation by gelatinase B depends on its localization to the cell surface by the hyaluronan receptor, CD44 (26). Activation by collagenase 3 might require similar localization, possibly by a different mechanism.

Data base searching identified potential collagenase 3 cleavage sites in type IV collagen and biglycan. Type IV collagen is the major structural protein of basement membranes. Degradation of type IV collagen could lead to breakdown of the basement membrane, which occurs in cancer (33) and multiple sclerosis (34), thus suggesting a possible role for collagenase 3 in these diseases. Biglycan is a small proteoglycan that has homology to decorin and is found in connective tissue (28). Degradation of biglycan could be important in joint disease (21).

In summary, we have used phage display technology to identify a number of new collagenase 3 substrate sequences. Many of the new substrate sequences are similar to known cleavage sites in type II collagen or aggrecan. However, others deviate significantly from previously known cleavage sequences. Considered together with previous work, the phage display results help define the overall specificity of collagenase 3. We have also used this knowledge of collagenase 3 specificity to speculatively identify other possible substrates. Our data base search results support previous studies demonstrating that gelatinase B can degrade TGF-β (26) and provide some evidence for the first time that collagenase 3 may also be involved in processing of the latency-associated peptide to activate this growth factor. Future studies will help to determine if any of the proteins identified by data base searches are actual substrates for collagenase 3.

Acknowledgments—We thank Carlos López-Otin for providing cDNA for human collagenase 3, Perry Brignola and Anne Hassell for providing collagenase 1 and stromelysin 1, and Matthew Smith and Marc Navre for providing random hexapeptide substrate phage library.

REFERENCES

1. Freije, J. M. P., Diez-Itza, I., Balbin, M., Sanchez, L. M., Blasco, R., Tolivia, J., and Lopez-Otin, C. (1994) *J. Biol. Chem.* **269,** 16766–16773
2. Reboul, P., Pelletier, J.-P., Tardif, G., Cloutier, J.-M., and Martel-Pelletier, J. (1996) *J. Clin. Invest.* **97,** 2011–2019
3. Billinghurst, R. C., Dahlberg, L., Ionescu, M., Reiner, A., Bourne, R., Rorabeck, C., Mitchell, P., Hambor, J., Diekmann, O., Tschesche, H., Chen, J., Wart, H. V., and Poole, A. R. (1997) *J. Clin. Invest.* **99,** 1534–1545
4. Willmroth, F., Peter, H. H., and Conca, W. (1998) *Immunobiology* **198,** 375–384
5. Mitchell, P. G., Magna, H. A., Reeves, L. M., Lopresti-Morrow, L. L., Yocum, S. A., Rosner, P. J., Geoghegan, K. F., and Hambor, J. E. (1996) *J. Clin. Invest.* **97,** 761–768
6. Knaüper, V., López-Otin, C., Smith, B., Knight, G., and Murphy, G. (1996) *J. Biol. Chem.* **271,** 1544–1550
7. Fosang, A. J., Last, K., Knauper, V., Murphy, G., and Neame, P. J. (1996) *FEBS Lett.* **380,** 17–20
8. Knaüper, V., Cowell, S., Smith, B., López-Otin, C., O'Shea, M., Morris, H., Zardi, L., and Murphy, G. (1997) *J. Biol. Chem.* **272,** 7608–7616
9. Kassel, D. B., Green, M. D., Wehbie, R. S., Swanstrom, R., and Berman, J. (1995) *Anal. Biochem.* **228,** 259–266
10. Tozser, J., Bagossi, P., Weber, I. T., Copeland, T. D., and Oroszlan, S. (1996) *J. Biol. Chem.* **271,** 6781–6788
11. Berman, J., Green, M., Sugg, E., Anderegg, R., Millington, D. S., Norwood, D. L., McGeehan, J., and Wiseman, J. (1992) *J. Biol. Chem.* **267,** 1434–1437
12. McGeehan, G. M., Bickett, D. M., Green, M., Kassel, D., Wiseman, J. S., and Berman, J. (1994) *J. Biol. Chem.* **269,** 32814–32820
13. Singh, J., Allen, M. P., Ator, M. A., Gainor, J. A., Whipple, D. A., Solowiej, J. E., Treasurywala, A. M., and Upson, D. A. (1995) *J. Med. Chem.* **38,** 217–219
14. Matthews, D. J., and Wells, J. A. (1993) *Science* **260,** 1113–1117
15. Smith, M. M., Shi, L., and Navre, M. (1995) *J. Biol. Chem.* **270,** 6440–6449
16. Ding, L., Coombs, G. S., Strandberg, L., Navre, M., Corey, D. R., and Madison, E. L. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 7627–7631
17. Stack, M. S., and Gray, R. D. (1989) *J. Biol. Chem.* **264,** 4277–4281
18. Knight, C. G., Willenbrock, F., and Murphy, G. (1992) *FEBS Lett.* **296,** 263–266
19. McGeehan, G. M., Becherer, J. D., Bast, R. C. J., Boyer, C. M., Champion, B., Connolly, K. M., Conway, J. G., Furdon, P., Karp, S., Kidao, S., McElroy, A. B., Nichols, J., Pryzwansky, K. M., Schoenen, F., Sekut, L, Truesdale, A., Verghese, M., Warner, J., and Ways, J. P. (1994) *Nature* **370,** 558–561
20. Hassell, A. M., Anderegg, R. J., Weigl, D., Milburn, M. V., Burkhart, W., Smith, G. F. H., Graber, P., Wells, T. N. C., Luther, M. A., and Jordan, S. R. (1994) *J. Mol. Biol.* **236,** 1410–1412
21. Hardingham, T. E., and Fosang, A. J. (1992) *FASEB J.* **6,** 861–870
22. Lauer-Fields, J. L., Tuzinski, K. A., Shimokawa, K., Nagase, H., and Fields, G. B. (2000) *J. Biol. Chem.* **275,** 13282–13290
23. Nagase, H., Fields, C. G., and Fields, G. B. (1994) *J. Biol. Chem.* **269,** 20952–20957
24. Nagase, H., and Fields, G. B. (1996) *Biopolymers* **40,** 399–416
25. Beekman, B., van El, B., Drijfhout, J. W., Ronday, H. K., and TeKoppele, J. M. (1997) *FEBS Lett.* **418,** 305–309
26. Yu, Q., and Stamenkovic, I. (2000) *Genes Dev.* **14,** 163–176
27. Fields, G. B.(1991) *J. Theor. Biol.* **153,** 585–602
28. Fisher, L. W., Termine, J. D., and Young, M. F. (1989) *J. Biol. Chem.* **264,** 4571–4576
29. Lovejoy, B., Welch, A. R., Carr, S., Luong, C., Broka, C., Hendricks, R. T., Campbell, J. A., Walker, K. A. M., Martin, R., Van Wart, H., and Browner, M. F. (1999) *Nat. Struct. Biol.* **6,** 217–221
30. Cox, D. A. (1995) *Cell Biol. Int.* **18,** 357–371
31. Miyazono, K., and Heldin, C. H. (1991) *CIBA Found. Symp.* **157,** 81–92
32. Flaumenhaft, R., Kojima, S., Abe, M., and Rifkin, D. B. (1993) *Adv. Pharmacol.* **24,** 51–76
33. Shapiro, S. D. (1998) *Curr. Opin. Cell Biol.* **10,** 602–608
34. Parish, C. R., Hindmarsh, E. J., Bartlett, M. R., Staykova, M. A., Cowden, W. B., and Willenborg, D. O. (1998) *Immunol. Cell Biol.* **76,** 104–113