ATTACHMENT F

U.S. Patent No. 6,054,297
References Cited

U.S. PATENT DOCUMENTS

4,845,198 7/1989 Urdal et al.
5,132,405 7/1992 Huston et al.
5,225,539 7/1993 Winter.
5,530,101 6/1996 Queen et al.
5,558,864 9/1996 Bendig et al.
5,585,080 12/1996 Queen et al. .......... 424/133.1
5,677,171 10/1997 Hoobliak et al.
5,693,761 12/1997 Queen et al.
5,693,762 12/1997 Queen et al.
5,714,350 2/1998 Cot et al.
5,859,205 1/1999 Adair et al.

FOREIGN PATENT DOCUMENTS

8505891 3/1992 Australia
126094 10/1984 European Pat. Off.
1,950274 A1 11/1984 European Pat. Off.

OTHER PUBLICATIONS

Queen et al., PNAS, USA, 85:10029-10033, 1989.

(List continued on next page.)

Primary Examiner—Paula K. Hutzeil
Assistant Examiner—Julie E. Reyes
Attorney, Agent, or Firm—Wendy M. Lee

ABSTRACT

Variant immunoglobulins, particularly humanized antibody polypeptides are provided, along with methods for their preparation and use. Consensus immunoglobulin sequences and structural models are also provided.

30 Claims, 12 Drawing Sheets
OTHER PUBLICATIONS


Tao et al., "Role of Carbohydrate in the Structure and Effector Functions Meditated by the Human IgG Constant Region" *Journal of Immunology* 143(8):2595-2601 (1989).


Orlandi et al., "Cloning of cDNA Corresponding to Heavy and Light Chain Immunoglobulin Variable Domains" *Protein and Pharmaceutical Engineering* p. 90 (1989).


Riechmann & Winter, "Recombinant Antibodies" (II. of London Royal Postgraduate Medical School, Wolfson Institute, Abstracts) (May 1987).


Shields et al., "Inhibition of Allergic Reactions with Antibodies to IgE" International Archives of Allergy and Immunology 107 (1-3):308–312 (May 1995).


Tighe et al., "Delayed Allograft Rejection in Primates Treated with Anti-IL-2 Receptor Monoclonal antibody Campath-6" Transplantation 45(1):226–228 (Jan. 1988).


**FIG. 1A**

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>4D5</td>
<td>DIVMTQS</td>
<td>HU4D5</td>
<td>DIQMTQSP</td>
<td>HUV&lt;sub&gt;L&lt;/sub&gt;</td>
<td>DIQMTQSP</td>
</tr>
<tr>
<td></td>
<td>SHKFMTST</td>
<td>SSSLASVG</td>
<td>SSLASVG</td>
<td>SSYS</td>
<td>SSLASVG</td>
</tr>
<tr>
<td></td>
<td>SVGDRVST</td>
<td>DTITCRASQ</td>
<td>DTITCRASQ</td>
<td>DTITCRASQ</td>
<td>DTITCRASQ</td>
</tr>
<tr>
<td></td>
<td>ITCKASQDV</td>
<td>NVTAVA</td>
<td>QVYQQ</td>
<td>WP</td>
<td>NVTAVA</td>
</tr>
<tr>
<td></td>
<td>HPGSP</td>
<td>K</td>
<td>QQ</td>
<td>KPG</td>
<td>Q</td>
</tr>
<tr>
<td></td>
<td>KLIYSASF</td>
<td>FLES</td>
<td>FLES</td>
<td>FLES</td>
<td>FLES</td>
</tr>
</tbody>
</table>

---

V<sub>L</sub>-CDR1

---

V<sub>L</sub>-CDR2

---

<table>
<thead>
<tr>
<th></th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>4D5</td>
<td>GVPDRF</td>
<td>HU4D5</td>
<td>GVPSRF</td>
<td>HUV&lt;sub&gt;L&lt;/sub&gt;</td>
<td>GVPSRF</td>
</tr>
<tr>
<td></td>
<td>TGNRS</td>
<td>GGGSSG</td>
<td>GTDGRTL</td>
<td>GGGSSG</td>
<td>GTDGRTL</td>
</tr>
<tr>
<td></td>
<td>GSDFTFTTISSVQAEDLA</td>
<td>TISSLQPED</td>
<td>FATYYCQHY</td>
<td>TISSLQPED</td>
<td>FATYYCQHY</td>
</tr>
<tr>
<td></td>
<td>VYCCQHY</td>
<td>TP</td>
<td>GQ</td>
<td>TP</td>
<td>GQ</td>
</tr>
<tr>
<td></td>
<td>TPF</td>
<td>GGT</td>
<td>KLEIKRA</td>
<td>GGT</td>
<td>KLEIKRA</td>
</tr>
<tr>
<td></td>
<td>TTPTFGG</td>
<td>GTKVEIKRT</td>
<td>TTPTFGG</td>
<td>GTKVEIKRT</td>
<td>TTPTFGG</td>
</tr>
</tbody>
</table>

---

V<sub>L</sub>-CDR3
Anneal huV<sub>L</sub> or huV<sub>H</sub> oligomers to pAK1 template

1. Ligate
2. Isolate assembled oligomers
3. Anneal to pAK1 template (Xhol<sup>-</sup>, SstI<sup>+</sup>)
4. Extend and ligate

1. Transform <i>E. coli</i>
2. Isolate phagemid pool
3. Enrich for huV<sub>L</sub> and huV<sub>H</sub>(Xho I<sup>+</sup>, SstI<sup>-</sup>)
4. Sequence verify

FIG. 2
FIG. 3
FIG. 5
**FIG. 6A-1**

```
H52H4-160  FIG. 6A-1  QVQLQSGPELVKPGASVKISCKTSGYTFTE
         10       20      30
pH52-8.0  MGWSCIILFLVATATGVHEQVLESGGLVQPGGSLRSLSCATEGYTFTE
         10       20      30
                  40      50      60      70      80
H52H4-160  YTMHWKQSHGKSLWIGGFNPKNGSSHNRQFMKATLAVDVKSTAYM
         90      100     110     120     130
pH52-8.0  YTMHWMRQAPGKLEWVGAINPKNGGTSHQRFMDRTISVDKSTAYM
         60       70       80       90      100
                  110     120     130     140     150
H52H4-160  ELRSLETSEDGIYYCARWGLGYFDFVYFDVGAFTTVSASTKPS
         140     150     160     170     180
pH52-8.0  QUMLRAEDTAVYCARWGLGYFDFVYFDVQGQTVSASTKPS
         110     120     130     140     150
                  160     170     180     190     200
H52H4-160  VFPLAPSSKSTSGTAAAGCLVYDFEPRTVSNSGALTSGVHTFPAVL
         190     200     210     220     230
pH52-8.0  VFPLAPCRSTSESTAAAGCLVYDFEPRTVSNSGALTSGVHTFPAVL
         160     170     180     190     200
                  210     220     230     240
H52H4-160  QSSGLYSLSSVTVPSLGLGTQYICNVNHKPSMTKVDKVPEKSCDKTH
         240     250     260     270     280
pH52-8.0  QSSGLYSLSSVTVTSNFTQYTCNDHFKPSMTKVDKTVKCC---V
         210     220     230     240
                  250     260     270     280     290
H52H4-160  TCPPCAPPSELLGGSVFLFP PKKDTLMISRTPEVTFCVVDVSHEDPEVK
         240     250     260     270     280
pH52-8.0  ECPCAPAPPL-GAPSVFLFP PKDLMISRTPEVTFCVVDVSHEDPEVQ
         250     260     270     280     290
```
**FIG. 6A-2**

<table>
<thead>
<tr>
<th></th>
<th>290</th>
<th>300</th>
<th>310</th>
<th>320</th>
<th>330</th>
</tr>
</thead>
<tbody>
<tr>
<td>H52H4-160</td>
<td>FNWYV</td>
<td>DGVE</td>
<td>VHNA</td>
<td>KTKP</td>
<td>REEQ</td>
</tr>
<tr>
<td>pH52-8.0</td>
<td>FYST</td>
<td>RNVS</td>
<td>LTVL</td>
<td>HVGQ</td>
<td>DWLNGKEYKCKVS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>340</th>
<th>350</th>
<th>360</th>
<th>370</th>
<th>380</th>
</tr>
</thead>
<tbody>
<tr>
<td>H52H4-160</td>
<td>NKA</td>
<td>LPAP</td>
<td>IEKT</td>
<td>ISK</td>
<td>AKG</td>
</tr>
<tr>
<td>pH52-8.0</td>
<td><strong>,</strong></td>
<td>*********</td>
<td>**********</td>
<td>**********</td>
<td>**********</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>390</th>
<th>400</th>
<th>410</th>
<th>420</th>
<th>430</th>
</tr>
</thead>
<tbody>
<tr>
<td>H52H4-160</td>
<td>SDIA</td>
<td>VE</td>
<td>WESNGQPENNYKTTPV</td>
<td>LDSDG</td>
<td>FFLYSK</td>
</tr>
<tr>
<td>pH52-8.0</td>
<td>**********</td>
<td>**********</td>
<td>**********</td>
<td>**********</td>
<td>**********</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>440</th>
<th>450</th>
</tr>
</thead>
<tbody>
<tr>
<td>H52H4-160</td>
<td>CSV</td>
<td>VMHEALHN</td>
</tr>
<tr>
<td>pH52-8.0</td>
<td>**********</td>
<td>**********</td>
</tr>
</tbody>
</table>
**FIG. 6B**

<table>
<thead>
<tr>
<th>pH52-9.0</th>
<th>pH52-9.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>H52L6-158</td>
<td>H52L6-158</td>
</tr>
<tr>
<td><strong>10</strong></td>
<td><strong>10</strong></td>
</tr>
<tr>
<td>DVQMTQTSLGASLGDRVTINCRASQDNN</td>
<td>DVQMTQTSLGASLGDRVTINCRASQDNN</td>
</tr>
<tr>
<td><strong>20</strong></td>
<td><strong>20</strong></td>
</tr>
<tr>
<td>MGWSCIILVLATATGVHSDIQMTQSPSSLASVGRVT lgITCRASQDNN</td>
<td>MGWSCIILVLATATGVHSDIQMTQSPSSLASVGRVT lgITCRASQDNN</td>
</tr>
<tr>
<td><strong>30</strong></td>
<td><strong>30</strong></td>
</tr>
<tr>
<td><strong>40</strong></td>
<td><strong>40</strong></td>
</tr>
<tr>
<td><strong>50</strong></td>
<td><strong>50</strong></td>
</tr>
<tr>
<td><strong>60</strong></td>
<td><strong>60</strong></td>
</tr>
<tr>
<td><strong>70</strong></td>
<td><strong>70</strong></td>
</tr>
<tr>
<td><strong>80</strong></td>
<td><strong>80</strong></td>
</tr>
<tr>
<td><strong>90</strong></td>
<td><strong>90</strong></td>
</tr>
<tr>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td><strong>110</strong></td>
<td><strong>110</strong></td>
</tr>
<tr>
<td><strong>120</strong></td>
<td><strong>120</strong></td>
</tr>
<tr>
<td><strong>130</strong></td>
<td><strong>130</strong></td>
</tr>
<tr>
<td><strong>140</strong></td>
<td><strong>140</strong></td>
</tr>
<tr>
<td><strong>150</strong></td>
<td><strong>150</strong></td>
</tr>
<tr>
<td><strong>160</strong></td>
<td><strong>160</strong></td>
</tr>
<tr>
<td><strong>170</strong></td>
<td><strong>170</strong></td>
</tr>
<tr>
<td><strong>180</strong></td>
<td><strong>180</strong></td>
</tr>
<tr>
<td><strong>190</strong></td>
<td><strong>190</strong></td>
</tr>
<tr>
<td><strong>200</strong></td>
<td><strong>200</strong></td>
</tr>
<tr>
<td><strong>210</strong></td>
<td><strong>210</strong></td>
</tr>
<tr>
<td><strong>220</strong></td>
<td><strong>220</strong></td>
</tr>
<tr>
<td><strong>230</strong></td>
<td><strong>230</strong></td>
</tr>
<tr>
<td><strong>240</strong></td>
<td><strong>240</strong></td>
</tr>
<tr>
<td><strong>250</strong></td>
<td><strong>250</strong></td>
</tr>
<tr>
<td><strong>260</strong></td>
<td><strong>260</strong></td>
</tr>
<tr>
<td><strong>270</strong></td>
<td><strong>270</strong></td>
</tr>
<tr>
<td><strong>280</strong></td>
<td><strong>280</strong></td>
</tr>
<tr>
<td><strong>290</strong></td>
<td><strong>290</strong></td>
</tr>
<tr>
<td><strong>300</strong></td>
<td><strong>300</strong></td>
</tr>
<tr>
<td><strong>310</strong></td>
<td><strong>310</strong></td>
</tr>
<tr>
<td><strong>320</strong></td>
<td><strong>320</strong></td>
</tr>
<tr>
<td><strong>330</strong></td>
<td><strong>330</strong></td>
</tr>
<tr>
<td><strong>340</strong></td>
<td><strong>340</strong></td>
</tr>
<tr>
<td><strong>350</strong></td>
<td><strong>350</strong></td>
</tr>
<tr>
<td><strong>360</strong></td>
<td><strong>360</strong></td>
</tr>
<tr>
<td><strong>370</strong></td>
<td><strong>370</strong></td>
</tr>
<tr>
<td><strong>380</strong></td>
<td><strong>380</strong></td>
</tr>
<tr>
<td><strong>390</strong></td>
<td><strong>390</strong></td>
</tr>
<tr>
<td><strong>400</strong></td>
<td><strong>400</strong></td>
</tr>
<tr>
<td><strong>410</strong></td>
<td><strong>410</strong></td>
</tr>
<tr>
<td><strong>420</strong></td>
<td><strong>420</strong></td>
</tr>
<tr>
<td><strong>430</strong></td>
<td><strong>430</strong></td>
</tr>
<tr>
<td><strong>440</strong></td>
<td><strong>440</strong></td>
</tr>
<tr>
<td><strong>450</strong></td>
<td><strong>450</strong></td>
</tr>
<tr>
<td><strong>460</strong></td>
<td><strong>460</strong></td>
</tr>
<tr>
<td><strong>470</strong></td>
<td><strong>470</strong></td>
</tr>
<tr>
<td><strong>480</strong></td>
<td><strong>480</strong></td>
</tr>
<tr>
<td><strong>490</strong></td>
<td><strong>490</strong></td>
</tr>
<tr>
<td><strong>500</strong></td>
<td><strong>500</strong></td>
</tr>
<tr>
<td><strong>510</strong></td>
<td><strong>510</strong></td>
</tr>
<tr>
<td><strong>520</strong></td>
<td><strong>520</strong></td>
</tr>
<tr>
<td><strong>530</strong></td>
<td><strong>530</strong></td>
</tr>
<tr>
<td><strong>540</strong></td>
<td><strong>540</strong></td>
</tr>
<tr>
<td><strong>550</strong></td>
<td><strong>550</strong></td>
</tr>
<tr>
<td><strong>560</strong></td>
<td><strong>560</strong></td>
</tr>
<tr>
<td><strong>570</strong></td>
<td><strong>570</strong></td>
</tr>
<tr>
<td><strong>580</strong></td>
<td><strong>580</strong></td>
</tr>
<tr>
<td><strong>590</strong></td>
<td><strong>590</strong></td>
</tr>
<tr>
<td><strong>600</strong></td>
<td><strong>600</strong></td>
</tr>
<tr>
<td><strong>610</strong></td>
<td><strong>610</strong></td>
</tr>
<tr>
<td><strong>620</strong></td>
<td><strong>620</strong></td>
</tr>
<tr>
<td><strong>630</strong></td>
<td><strong>630</strong></td>
</tr>
<tr>
<td><strong>640</strong></td>
<td><strong>640</strong></td>
</tr>
<tr>
<td><strong>650</strong></td>
<td><strong>650</strong></td>
</tr>
<tr>
<td><strong>660</strong></td>
<td><strong>660</strong></td>
</tr>
<tr>
<td><strong>670</strong></td>
<td><strong>670</strong></td>
</tr>
<tr>
<td><strong>680</strong></td>
<td><strong>680</strong></td>
</tr>
<tr>
<td><strong>690</strong></td>
<td><strong>690</strong></td>
</tr>
<tr>
<td><strong>700</strong></td>
<td><strong>700</strong></td>
</tr>
<tr>
<td><strong>710</strong></td>
<td><strong>710</strong></td>
</tr>
<tr>
<td><strong>720</strong></td>
<td><strong>720</strong></td>
</tr>
<tr>
<td><strong>730</strong></td>
<td><strong>730</strong></td>
</tr>
<tr>
<td><strong>740</strong></td>
<td><strong>740</strong></td>
</tr>
<tr>
<td><strong>750</strong></td>
<td><strong>750</strong></td>
</tr>
<tr>
<td><strong>760</strong></td>
<td><strong>760</strong></td>
</tr>
<tr>
<td><strong>770</strong></td>
<td><strong>770</strong></td>
</tr>
<tr>
<td><strong>780</strong></td>
<td><strong>780</strong></td>
</tr>
<tr>
<td><strong>790</strong></td>
<td><strong>790</strong></td>
</tr>
<tr>
<td><strong>800</strong></td>
<td><strong>800</strong></td>
</tr>
<tr>
<td><strong>810</strong></td>
<td><strong>810</strong></td>
</tr>
<tr>
<td><strong>820</strong></td>
<td><strong>820</strong></td>
</tr>
<tr>
<td><strong>830</strong></td>
<td><strong>830</strong></td>
</tr>
<tr>
<td><strong>840</strong></td>
<td><strong>840</strong></td>
</tr>
<tr>
<td><strong>850</strong></td>
<td><strong>850</strong></td>
</tr>
<tr>
<td><strong>860</strong></td>
<td><strong>860</strong></td>
</tr>
<tr>
<td><strong>870</strong></td>
<td><strong>870</strong></td>
</tr>
<tr>
<td><strong>880</strong></td>
<td><strong>880</strong></td>
</tr>
<tr>
<td><strong>890</strong></td>
<td><strong>890</strong></td>
</tr>
<tr>
<td><strong>900</strong></td>
<td><strong>900</strong></td>
</tr>
<tr>
<td><strong>910</strong></td>
<td><strong>910</strong></td>
</tr>
<tr>
<td><strong>920</strong></td>
<td><strong>920</strong></td>
</tr>
<tr>
<td><strong>930</strong></td>
<td><strong>930</strong></td>
</tr>
<tr>
<td><strong>940</strong></td>
<td><strong>940</strong></td>
</tr>
<tr>
<td><strong>950</strong></td>
<td><strong>950</strong></td>
</tr>
<tr>
<td><strong>960</strong></td>
<td><strong>960</strong></td>
</tr>
<tr>
<td><strong>970</strong></td>
<td><strong>970</strong></td>
</tr>
<tr>
<td><strong>980</strong></td>
<td><strong>980</strong></td>
</tr>
<tr>
<td><strong>990</strong></td>
<td><strong>990</strong></td>
</tr>
<tr>
<td><strong>1000</strong></td>
<td><strong>1000</strong></td>
</tr>
</tbody>
</table>
FIG. 7A-1

| verA.hcfab  | 1 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verI.hcfab  | 1 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verN.hcfab  | 1 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verO.hcfab  | 1 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verO.hcfab2 | 5 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verP.hcfab  | 6 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verQ.hcfab  | 7 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verR.hcfab  | 8 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verS.hcfab  | 9 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verT.hcfab  | 10 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verU.hcfab  | 11 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verV.hcfab  | 12 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verW.hcfab  | 13 | EVQLVESGGGLVQPGSGSRLSCATSGYTFETYTMNWRQAPGKسلGlewag |
| verA.hcfab  | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verI.hcfab  | 51 | INFNGKTSNHQRPMDFTRPSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verN.hcfab  | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verO.hcfab  | 51 | INFNGKTSNHQRPMDFTRPSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verO.hcfab2 | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verP.hcfab  | 51 | INFNGKTSNHQRPMDFTRPSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verQ.hcfab  | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verR.hcfab  | 51 | INFNGKTSNHQRPMDFTRPSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verS.hcfab  | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verT.hcfab  | 51 | INFNGKTSNHQRPMDFTRPSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verU.hcfab  | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verV.hcfab  | 51 | INFNGKTSNHQRPMDFTRPSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verW.hcfab  | 51 | INFNGKTSYARVKPGRTSVDKOENTLYQMNLSLRAEDTAVYCARWR |
| verA.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verI.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verN.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verO.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verO.hcfab2 | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verP.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verQ.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verR.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verS.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verT.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verU.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verV.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |
| verW.hcfab  | 101 | GLNYGFDVRYFDVWQGQTTLVITSSASTKPSVFPLAPSSKSTSGTAALG |

SEQ.ID.NO.26

SEQ.ID.NO.27

SEQ.ID.NO.28

SEQ.ID.NO.29

SEQ.ID.NO.30

SEQ.ID.NO.31

SEQ.ID.NO.32

SEQ.ID.NO.33

SEQ.ID.NO.34

SEQ.ID.NO.35

SEQ.ID.NO.36

SEQ.ID.NO.37

SEQ.ID.NO.38
FIG. 7A-2

verA.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verI.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verN.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verO.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verO.hcfab2 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verP.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verQ.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verR.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verS.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verT.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verU.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verV.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL
verW.hcfab 151 CLVKDYFPEPVTQSWNSGALTSGVHTFPAVLQSSGLYSLSSSVTVTVPSSL

verA.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verI.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verN.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verO.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verO.hcfab2 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verP.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verQ.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verR.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verS.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verT.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verU.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verV.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
verW.hcfab 201 GTQTYICNVNHKPSNTKVDDKKEPKSCKTHT---
FIG. 7B

<table>
<thead>
<tr>
<th></th>
<th>verAlc</th>
<th>verZlc</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>DIQMTQSPSSLASAVGDRTITCRASTQDNYYLNYQQKPGKAPKLIIYY</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>DIQMTQSPSSLASAVGDRTITCRASTQDNYYLNYQQKPGKAPKLIIYY</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>TSTLBSVPSRSGSSTDYTITISSLQEDFATYPQNNLPPPTFGQ</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>TSTLHSVPSRSGSSTDYTITISSLQEDFATYPQNNLPPPTFGQ</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>GTKVBIKRTVAAPSVFIFPPSDEQLKSGTASVVCLNPPFYKVAKVQWKV</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>GTKVBIKRTVAAPSVFIFPPSDEQLKSGTASVVCLNPPFYKVAKVQWKV</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>DNALQSGNSQESVTQDSDKSTYSLSSTLSKADYKHKVIARVTQHG</td>
<td></td>
</tr>
<tr>
<td>170</td>
<td>DNALQSGNSQESVTQDSDKSTYSLSSTLSKADYKHKVIARVTQHG</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>LSSFVTQPSFRGEC</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>LSSFVTQPSFRGEC</td>
<td></td>
</tr>
</tbody>
</table>

1. (SEQ.ID.NO.39)
2. (SEQ.ID.NO.40)
HUMANIZED ANTIBODIES AND METHODS FOR MAKING THEM


FIELD OF THE INVENTION

This invention relates to methods for the preparation and use of variant antibodies and finds application particularly in the fields of immunology and cancer diagnosis and therapy.

BACKGROUND OF THE INVENTION

Naturally occurring antibodies (immunoglobulins) comprise two heavy chains linked together by disulfide bonds and two light chains, one light chain being linked to each of the heavy chains by disulfide bonds. Each heavy chain has at one end a variable domain (V H) followed by a number of constant domains. Each light chain has a variable domain (V L) at one end and a constant domain at its other end; the constant domain of the light chain is aligned with the first constant domain of the heavy chain, and the light chain variable domain is aligned with the variable domain of the heavy chain. Particular amino acid residues are believed to form an interface between the light and heavy chain variable domains, see e.g. Chothia et al., J. Mol. Biol. 180:651-663 (1985); Novotny and Haber, Proc. Natl. Acad. Sci. USA 82:4592-4596 (1985).

The constant domains are not involved directly in binding the antibody to an antigen, but are involved in various effector functions, such as participation of the antibody to antibody-dependent cellular cytotoxicity. The variable domains of each pair of light and heavy chains are involved directly in binding the antibody to the antigen. The domains of natural light and heavy chains have the same general structure, and each domain comprises four framework (FR) regions, whose sequences are somewhat conserved, connected by three hyper-variable or complementarity determining regions (CDRs) (see Kabat, E. A. et al., Sequences of Proteins of Immunological Interest, National Institutes of Health, Bethesda, Md. (1987)). The four framework regions largely adopt a beta-sheet conformation and the CDRs form loops connecting, and in some cases forming part of, the beta-sheet structure. The CDRs in each chain are held in close proximity by the framework regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site.

Widespread use has been made of monoclonal antibodies, particularly those derived from rodents including mice, however they are frequently antigenic in human clinical use. For example, a major limitation in the clinical use of rodent monoclonal antibodies is an anti-globulin response during therapy (Miller, R. A. et al., Blood 62:988-995 (1983); Schroff, R. W. et al., Cancer Res. 45:879-885 (1985)).

The art has attempted to overcome this problem by constructing "chimeric" antibodies in which an animal antibody binding variable domain is coupled to a human constant domain (Cabilly et al., U.S. Pat. No. 4,816,567; Morrison, S. L. et al., Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984); Boulianne, G. L. et al., Nature 312:643-646 (1984); Neuberger, M. S. et al., Nature 314:260-270 (1985)). The term "chimeric" antibody is used herein to describe a polypeptide comprising at least the antigen binding portion of an antibody molecule linked to at least part of another protein (typically an immunoglobulin constant domain).


In the typical embodiment, such chimeric antibodies contain about one third rodent (or other non-human species) sequence and thus are capable of eliciting a significant anti-globulin response in humans. For example, in the case of the murine anti-CD3 antibody, OKT3, much of the resulting anti-globulin response is directed against the variable region rather than the constant region (Jeffers, G. J. et al., Transplantation 41:572-578 (1986)).

In a further effort to resolve the antigen binding functions of antibodies and to minimize the use of heterologous sequences in human antibodies, Winter and colleagues (Jones, P. T. et al., Nature 292:225-228 (1986); Riechmann, L. et al., Nature 332:322-327 (1988); Verhoeyen, M. et al., Science 239:1534-1536 (1988)) have substituted rodent CDRs or CDR sequences for the corresponding segments of a human antibody. As used herein, the term "humanized" antibody is an embodiment of chimeric antibodies wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.


In some cases, substituting CDRs from rodent antibodies for the human CDRs in human frameworks is sufficient to transfer high antigen binding affinity (Jones, P. T. et al., Nature 321:522-525 (1986); Verhoeyen, M. et al., Science 239:1534-1536 (1988)), whereas in other cases it has been necessary to additionally replace one (Riechmann, L. et al., Nature 332:322-327 (1988)) or several (Queen, C. et al., Proc. Natl. Acad. Sci. USA 86:10029-10033 (1989)) framework region (FR) residues. See also Co et al., supra.

For a given antibody a small number of FR residues are anticipated to be important for antigen binding. Firstly for example, certain antibodies have been shown to contain a few FR residues which directly contact antigen in crystal structures of antibody-antigen complexes (e.g., reviewed in Davies, D. R. et al., Ann Rev. Biochem. 59:439-473 (1990)).

Humanizing an antibody with retention of high affinity for antigen and other desired biological activities is at present difficult to achieve using currently available procedures. Methods are needed for rationalizing the selection of sites for substitution in preparing such antibodies and thereby increasing the efficiency of antibody humanization.

The proto-oncogene HER2 (human epidermal growth factor receptor 2) encodes a protein tyrosine kinase (p185HER2) that is related to and somewhat homologous to the human epidermal growth factor receptor (see Coussens, L. et al., Science 230:1132-1139 (1983); Yamamoto, T. et al., Nature 319:220-224 (1986); King, C. R. et al., Science 229:974-976 (1985)). HER2 is also known in the field as c-erbB-2, and sometimes by the name of the rat homolog, neu. Amplification and/or overexpression of HER2 is associated with multiple human malignancies and appears to be integrally involved in progression of 25-30% of human breast and ovarian cancers (Slamon, D. J. et al., Science 235:177-182 (1987), Slamon, D. J. et al., Science 244:707-712 (1989)). Furthermore, the extent of amplification is inversely correlated with the observed median patient survival time (Slamon, supra, Science 1989).
g. For any non-homologous import antibody amino acid residue which is reasonably expected to have at least one of these effects, substituting that residue for the corresponding amino acid residue in the consensus antibody FR sequence.

Optionally, the method of this invention comprises the additional steps of determining if any non-homologous residues identified in step (e) are exposed on the surface of the domain or buried within it, and if the residue is exposed but has none of the effects identified in step (f), retaining the consensus residue.


In certain embodiments, the method of this invention comprises the additional steps of searching either or both of the import, non-human and the consensus variable domain sequences for glycosylation sites, determining if the glycosylation is reasonably expected to be important for the desired antigen binding and biological activity of the antibody (i.e., determining if the glycosylation site binds to antigen or changes a side chain of an amino acid residue that binds to antigen, or if the glycosylation enhances or weakens antigen binding, or is important for maintaining antibody affinity). If the import sequence bears the glycosylation site, it is preferred to substitute that site for the corresponding residues in the consensus human if the glycosylation site is reasonably expected to be important. If only the consensus sequence, and not the import, bears the glycosylation site, it is preferred to eliminate that glycosylation site or substitute therefor the corresponding amino acid residues from the import sequence.

Another embodiment of this invention comprises aligning import antibody and the consensus antibody FR sequences, identifying import antibody FR residues which are non-homologous with the aligned consensus FR sequence, and for each such non-homologous import antibody FR residue, determining if the corresponding consensus antibody residue represents a residue which is highly conserved across all species at that site, and if it is so conserved, preparing a humanized antibody which comprises the corresponding amino acid residue at that site.

Certain alternate embodiments of the methods of this invention comprise determining the amino acid sequence of at least a portion of an import, non-human antibody variable domain having a CDR and a FR, obtaining the amino acid sequence of at least a portion of a consensus antibody variable domain having a CDR and a FR, substituting the non-human CDR for the human CDR in the consensus antibody variable domain, and then substituting an amino acid residue for the consensus amino acid residue at least one of the following sites:


In preferred embodiments, the non-CDR residue substituted at the consensus FR site is the residue found at the corresponding location of the non-human antibody.

Optionally, this just-recited embodiment comprises the additional steps of following the method steps appearing at the beginning of this summary and determining whether a particular amino acid residue can reasonably be expected to have undesirable effects.

This invention also relates to a humanized antibody comprising the CDR sequence of an import, non-human antibody and the FR sequence of a human antibody, wherein an amino acid residue within the human FR sequence located at any one of the sites 4L, 35L, 36L, 38L, 43L, 44L, 46L, 58L, 62L, 63L, 64L, 65L, 66L, 67L, 68L, 69L, 70L, 71L, 73L, 85L, 87L, 98L, 2H, 4H, 24H, 36H, 37H, 39H, 43H, 49H, 58H, 60H, 67H, 68H, 69H, 70H, 73H, 74H, 75H, 76H, 78H, 91H, 92H, 93H, and 103H has been substituted by another residue. In preferred embodiments, the residue substituted at the human FR site is the residue found at the corresponding location of the non-human antibody from which the non-human CDR was obtained. In other embodiments, no human FR residue other than those set forth in this group has been substituted.

This invention also encompasses specific humanized antibody variable domains, and isolated polypeptides having homology with the following sequences.

1. **SEQ. ID NO. 1**, which is the light chain variable domain of a humanized version of mUMAb4D5:

   ```
   D14QTV9PE6LSAQWGVYTVTCTACQSTAVAWQKQCPFQELVYA
   SFLSGPVVFRCSRGTPSTELIESLQPYEFATYYCQQYNSLPPM
   SFLESGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTP
   ```

2. **SEQ. ID NO. 2**, which is the heavy chain variable domain of a humanized version of mUMAb4D5:

   ```
   EFLVQGEGSGSLTPQPSLRLCSGQGRPLPETVIVNHPVRCA
   YPTNNQTVADDVGKEFPTISADTSDKATLQMNLSLAEETAVFCSMCR
   DFPMVYKQQLTVYVSS
   ```

In another aspect, this invention provides a consensus antibody variable domain amino acid sequence for use in the preparation of humanized antibodies, methods for obtaining, using, and storing a computer representation of such a consensus sequence, and computers comprising the sequence data of such a sequence. In one embodiment, the following consensus antibody variable domain amino acid sequences are provided:

**SEQ. ID NO. 3** (light chain):

```
D14QTV9PE6LSAQWGVYTVTCTACQSTAVAWQKQCPFQELVYA
SFLSGPVVFRCSRGTPSTELIESLQPYEFATYYCQQYNSLPPM
```

and

```
BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A shows the comparison of the V, domain amino acid residues of muMAb4D5, huMAb4D5, and a consensus sequence (SEQ. ID. NO. 1 and SEQ. ID. NO. 3, respectively). FIG. 1B shows the comparison between the V, domain amino acid residues of the huMAb4D5, huMAb4D5, and a consensus sequence (FIG. 1B, SEQ. ID. NO. 6, SEQ. ID. NO. 2 and SEQ. ID. NO. 4, respectively). Both FIGS. 1A and 1B use the generally accepted numbering scheme from Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md. (1987)). In both FIG. 1A and FIG. 1B, the CDR residues determined according to a standard sequence definition (as in Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) are indicated by the first underlining beneath the sequences, and the CDR residues determined according to a structural definition (as in Chothia, C. & Lesk, A. M., J. Mol. Biol. 196:901-917 (1987)) are indicated by the second, lower underlines. The mismatches between genes are shown by the vertical lines.

FIG. 2 shows a scheme for humanization of muMAb4D5 V, and V, by gene conversion mutagenesis.

FIG. 3 shows the inhibition of SK-BR-3 proliferation by MAb4D5 variants. Relative cell proliferation was determined as described (Hudziak, R. M. et al., .woZec. CeN. Viol. 9:1165-1172 (1989)) and data (average of triplicate determinations) are presented as a percentage of results with untreated cultures for muMAb4D5 (I), huMAb4D5-8 (n) and huMAb4D5-1 (P).

FIG. 4 shows a stereo view of a-carbon tracing for a model of huMAb4D5-8 V, and V, The CDR residues (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) are shown in bold and side chains of V, residues AT1, T73, A78, S93, Y102 and V, residues Y55 plus R66 (see Table 3) are shown.

FIG. 5 shows an amino acid sequence comparison of V, (top panel) and V, (lower panel) domains of the murine anti-CD18 monoclonal Ab UCHT1 (muxCD3, Shabat et al., J. Exp. Med. 175, 217-225 (1992) with a humanized variant of this antibody (huxCD3). Also shown are consensus sequences (most commonly occurring residue or pair of residues) of the most abundant human subgroups, namely V, cl and V, cl upon which the humanized sequences are based (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest, 6th edition, National Institutes of Health, Bethesda, Md., U.S.A. (1991)). The light chain sequences—muxCD3, huCD3v1 and huIII—correspond to SEQ. ID. NOs 16, 17, and 18, respectively. The heavy chain sequences—muxCD3, huCD3v1 and hu—correspond to SEQ. ID. NOs 19, 45, and 21, respectively. Residues which differ between muxCD3 and huCD3v1 are identified by an asterisk (*), whereas those which differ between humanized and consensus sequences are identified by a sharp sign (#). A bullet (•) denotes that a residue at this position has been found to contact antigen in one or more crystallographic structures of antibody/antigen complexes (Kabat et al., 1991; Mian, I. S. et al., J. Mol. Biol. 217, 133-151 (1991)). The location of CDR residues according to a sequence definition (Kabat et al., 1991) and a structural definition (Chothia and Lesk, suprap 1987) are shown by a line and carats (A) beneath the sequences, respectively.

FIGS. 6A-1 and 6A-2 compare murine and humanized amino acid sequences for the heavy chain of an anti-CD18 antibody. H52H4-160 (SEQ. ID. NO. 22) is the murine sequence, and pH52-8.0 (SEQ. ID. NO. 23) is the humanized heavy chain sequence. pH52-8.0 residue 143S is the final amino acid in the variable heavy domain V, (SEQ. ID. NO.46) and residue 144A is the first amino acid in the constant heavy domain C,.

FIG. 6B compares murine and humanized amino acid sequences for the light chain of an anti-CD18 antibody. H52L6-158 (SEQ. ID. NO. 24) is the murine sequence, and pH52-9.0 (SEQ. ID. NO. 25) is the humanized light chain sequence. pH52-9.0 residue 128T is the final amino acid in the light chain variable domain V, (SEQ. ID. NO.47), and residue 129V is the first amino acid in the light chain constant domain C,.

FIGS. 7A-1 and 7A-2 show an amino acid sequence alignment of the sequences of the heavy chains of thirteen humanized anti-CD18 (H52) variants (SEQ. ID. NOs 26-38).

FIG. 7B shows an amino acid sequence alignment of two humanized anti-CD18 (H52) light chain variants (SEQ. ID. NOs 39-40).

DETAILED DESCRIPTION OF THE INVENTION

Definitions

In general, the following words or phrases have the indicated definitions when used in the description, examples, and claims:

The murine monoclonal antibody known as huMAb4D5 (Fendly, B. M. et al., Cancer Res. 50:1551-1558 (1990)) is directed against the extracellular domain (ECD) of p185aaa-p. The muMAb4D5 and its uses are described in PCT application WO 89/06692 published Jul. 27, 1989. This murine antibody was deposited with the ATCC and designated ATCC CRL 10463. In this description and claims, the terms muMAb4D5, chMAb4D5 and huMAb4D5 represent murine, chimerized and humanized versions of the monoclonal antibody 4D5, respectively.

A humanized antibody for the purposes herein is an immunoglobulin amino acid sequence variant or fragment thereof which is capable of binding to a predetermined antigen and which comprises a FR region having substantially the amino acid sequence of a human immunoglobulin and a CDR having substantially the amino acid sequence of a non-human immunoglobulin.

Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are referred to herein as "import" residues, which are typically taken from an "import" antibody domain, particularly a variable domain. An import residue, sequence, or antibody has a desired affinity and/or specificity, or other desirable antibody biological activity as described herein.

In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains (Fab, Fab', F(ab')2, Fabc, Fv) in which all or
substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. Ordinarily, the antibody will contain both the light chain as well as at least the variable domain of a heavy chain. The antibody may also include the CH1, hinge, CH2, CH3, and CH4 regions of the heavy chain.

The humanized antibody will be selected from any class of immunoglobulins, including IgM, IgG, IgD, IgA and IgE, and any isotype, including IgG1, IgG2, IgG3 and IgG4. Usually the constant domain is a complement fixing constant domain where it is desired that the humanized antibody exhibit cytotoxic activity, and the class is typically IgG1. Where such cytotoxic activity is not desirable, the constant domain may be of the IgG2 class. The humanized antibody may comprise sequences from more than one class or isotype, and selecting particular constant domains to optimize desired effector functions is within the ordinary skill in the art.

The FR and CDR regions of the humanized antibody need not correspond precisely to the parental sequences, e.g., the import CDR or the consensus FR may be mutated by substitution, insertion or deletion of at least one residue so that the CDR or FR residue at that site does not correspond to either the consensus or the import antibody. Such mutations, however, will not be extensive. Usually, at least 75% of the humanized antibody residues will correspond to those of the parental FR and CDR sequences, more often 90%, and most preferably greater than 95%.

In general, humanized antibodies prepared by the method of this invention are produced by a process of analysis of the parental sequences and various conceptual humanized products using three dimensional models of the parental and humanized sequences. Three dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, i.e., the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen.

Residues that influence antigen binding are defined to be residues that are substantially responsible for the antigen affinity or antigen specificity of a candidate immunoglobulin, in a positive or a negative sense. The invention is directed to the selection and combination of FR residues from the consensus and import sequence so that the desired immunoglobulin characteristic is achieved. Such desired characteristics include increases in affinity and greater specificity for the target antigen, although it is conceivable that in some circumstances the opposite effects might be desired. In general, the CDR residues are directly and most substantially involved in influencing antigen binding (although not all CDR residues are so involved and therefore need not be substituted into the consensus sequence). However, FR residues also have a significant effect and can exert their influence in at least three ways: They may noncovalently directly bind to antigen, they may interact with CDR residues and they may affect the interface between the heavy and light chains.

A residue that noncovalently directly binds to antigen is one that, by three dimensional analysis, is reasonably expected to noncovalently directly bind to antigen. Typically, it is necessary to impute the position of antigen from the spatial location of neighboring CDRs and the dimensions and structure of the target antigen. In general, only those humanized antibody residues that are capable of forming salt bridges, hydrogen bonds, or hydrophobic interactions are likely to be involved in non-covalent antigen binding, however residues which have atoms which are separated from antigen spatially by 3.2 Angstroms or less may also non-covalently interact with antigen. Such residues typically are the relatively larger amino acids having the side chains with the greatest bulk, such as tyrosine, arginine, and lysine. Antigen-binding FR residues also typically will have side chains that are oriented into an envelope surrounding the solvent oriented face of a CDR which extends about 7 Angstroms into the solvent from the CDR domain and about 7 Angstroms on either side of the CDR domain, again as visualized by three dimensional modeling.

A residue that interacts with a CDR generally is a residue that either affects the conformation of the CDR polypeptide backbone or forms a noncovalent bond with a CDR residue side chain. Conformation-affecting residues ordinarily are those that change the spatial position of any CDR backbone atom (N, Cα, C, O, Cβ) by more than about 0.2 Angstroms. Backbone atoms of CDR sequences are displayed for example by residues that interrupt or modify organized structures such as beta sheets, helices or loops. Residues that can exert a profound affect on the conformation of neighboring sequences include proline and glycine, both of which can be induced to bend the backbone. Other residues that can displace backbone atoms are those that are capable of participating in salt bridges and hydrogen bonds.

A residue that interacts with a CDR side chain is one that is reasonably expected to form a noncovalent bond with a CDR side chain, generally either a salt bridge or hydrogen bond. Such residues are identified by three dimensional positioning of their side chains. A salt or ion bridge could be expected to form between two side chains positioned within about 2.5-3.2 Angstroms of one another that bear opposite charges, for example a lysinyl and a glutamyl pairing. A hydrogen bond could be expected to form between the side chains of residue pairs such as seryl or threonyl with aspartyl or glutamyl (or other hydrogen accepting residues). Such pairings are well known in the protein chemistry art and will be apparent to the artisan upon three dimensional modeling of the candidate immunoglobulin.

Immunoglobulin residues that affect the interface between heavy and light chain variable regions ("the VH - VL interface") are those that affect the proximity or orientation of the two chains with respect to one another. Certain residues involved in interchain interactions are already known and include Vh residues 34, 36, 38, 44, 46, 87, 89, 91, 96, and 98 and Vl residues 35, 37, 39, 45, 47, 91, 93, 95, 100, and 103 (utilizing the nomenclature set forth in Kabat et al., "Sequences of Proteins of Immunological Interest" (National Institutes of Health, Bethesda, Md., 1987)). Additional residues are newly identified by the inventors herein, and include 43L, 85L, 43H and 60H. While these residues are indicated for IgG only, they are applicable across species. In the practice of this invention, import antibody residues that are reasonably expected to be involved in interchain interactions are selected for substitution into the consensus sequence. It is believed that heretofore no humanized antibody has been prepared with an intrachain affecting residue selected from an import antibody sequence.

Since it is not entirely possible to predict in advance what the exact impact of a given substitution will be it may be
necessary to make the substitution and assay the candidate antibody for the desired characteristic. These steps, however, are per se routine and well within the ordinary skill of the art.

CDR and FR residues are determined according to a standard sequence definition (Kabat et al., *Sequences of Proteins of Immunological Interest*, National Institutes of Health, Bethesda Md. (1987), and a structural definition (as in Chothia and Lesk, *J Mol Biol.* 196:901-917 (1987)). Where these two methods result in slightly different identifications of a CDR, the structural definition is preferred, but the residues identified by the sequence definition method are considered important FR residues for determination of which framework residues to import into a consensus sequence.

Throughout this description, reference is made to the numbering scheme from Kabat, E. A., et al., *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Bethesda, Md. (1987) and (1991). In these compendiums, Kabat lists many amino acid sequences for antibodies for each subclass, and lists the most commonly occurring amino acid for each residue position in that subclass. Kabat uses a method for assigning a residue number to each amino acid in a listed sequence, and this method for assigning residue numbers has become standard in the field. The Kabat numbering scheme is followed in this description. For purposes of this invention, to assign residue numbers to a candidate antibody amino acid sequence which is not included in the Kabat compendium, one follows the following steps. Generally, the candidate sequence is aligned with any immunoglobulin sequence or any consensus sequence in Kabat. Alignment may be done by hand, or by computer using commonly accepted computer programs; an example of such a program is the Align 2 program discussed in this description. Alignment may be facilitated by using some amino acid residues which are common to most Fab sequences. For example, the light and heavy chains each typically have two cysteines which have the same residue numbers, in V domain the two cysteines are typically at residue numbers 22 and 92, and in the V_{\gamma} domain the two cysteine residues are typically numbered 22 and 92. Framework residues generally, but not always, have approximately the same number of residues, however the CDRs will vary in size. For example, in the case of a CDR from a candidate sequence which is longer than the CDR in the sequence in Kabat, to which it is aligned, typically suffixes are added to the residue number to indicate the insertion of additional residues (see, e.g. residues 100abcde in FIG. 5). For candidate sequences which, for example, align with a Kabat sequence for residues 34 and 36 but have no residue between them to align with residue 35, the number 35 is simply not assigned to a residue.

Thus, in humanization of an import variable sequence, one cuts out an entire human or consensus CDR and replaces it with an import CDR sequence, (a) the exact number of residues may be swapped, leaving the numbering the same, (b) fewer import amino acid residues may be introduced than are cut, in which case there will be a gap in the residue numbers, or (c) a larger number of amino acid residues may be introduced then were cut, in which case the numbering will involve the use of suffixes such as 100abcde. The terms “consensus sequence” and “consensus antibody” as used herein refers to an amino acid sequence which comprises the most frequently occurring amino acid residues at each location in all immunoglobulins of any particular subclass or subunit structure. The consensus sequence may be based on immunoglobulins of a particular species or of many species. A “consensus” sequence, structure, or antibody is understood to encompass a consensus human sequence as described in certain embodiments of this invention, and to refer to an amino acid sequence which comprises the most frequently occurring amino acid residues at each location in all human immunoglobulins of any particular subclass or subunit structure. This invention provides consensus human structures and consensus structures which consider other species in addition to human.

The subunit structures of the five immunoglobulin classes in humans are as follows:

<table>
<thead>
<tr>
<th>Class</th>
<th>Heavy Chain Subclasses</th>
<th>Light Chain Molecular Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgG</td>
<td>( \gamma_{1}, \gamma_{2}, \gamma_{3}, \gamma_{4} )</td>
<td>( \kappa ) or ( \lambda )</td>
</tr>
<tr>
<td>IgA</td>
<td>( \alpha )</td>
<td>( \kappa ) or ( \lambda )</td>
</tr>
<tr>
<td>IgM</td>
<td>( \mu )</td>
<td>( \kappa ) or ( \lambda )</td>
</tr>
<tr>
<td>IgE</td>
<td>( \epsilon )</td>
<td>( \kappa ) or ( \lambda )</td>
</tr>
</tbody>
</table>

In preferred embodiments of an IgG\( \gamma_{1} \) human consensus sequence, the consensus variable domain sequences are derived from the most abundant subclasses in the sequence compilation of Kabat et al., *Sequences of Proteins of Immunological Interest*, National Institutes of Health, Bethesda Md. (1987), namely V\( _{\gamma} \) subgroups I and V\( _{\gamma} \) group III. In such preferred embodiments, the V\( _{\gamma} \) consensus domain has the amino acid sequence:

\[
\text{SEQ: ID NO. } 3; \text{ EVQLVEEDGGLVQPSMEATIELVSLGAGASGGSSQVTLLAASNKGSSPAATNGSF} \\
\text{EVQQLVQPSDEADSGRTTSPKTITWYQFKSPGKMSTLLGTVVWNSDIYADSVKGFITSGSDEDSYVEHTQSSPEYFTKYLDGKVKSSEKSEDLTYYC}}
\]

These sequences include consensus CDRs as well as consensus FR residues (see for example in FIG 1).

While not wishing to be limited to any particular theories, it may be that these preferred embodiments are less likely to be immunogenic in an individual than less abundant subclasses. However, in other embodiments, the consensus sequence is derived from other subclasses of human immunoglobulin variable domains. In yet other embodiments, the consensus sequence is derived from human constant domains.

Identity or homology with respect to a specified amino acid sequence of this invention is defined herein as the percentage of amino acid residues in a candidate sequence that are identical with the specified residues, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity. None of N-terminal, C-terminal or internal extensions, deletions, or insertions into the specified sequence shall be constructed as affecting homology. All sequence alignments called for in this invention are based on conventional methods, a suitable computer program is the “Align 2” program for which protection is being sought from the U.S. Register of Copyrights (Align 2, by Genentech, Inc., application filed Dec. 9, 1991).
"Non-homologous" import antibody residues are those residues which are not identical to the amino acid residue at the analogous or corresponding location in a consensus sequence, after the import and consensus sequences are aligned.

The term "computer representation" refers to information which is in a form that can be manipulated by a computer. The act of storing a computer representation refers to the act of placing the information in a form suitable for manipulation by a computer.

This invention is also directed to novel polypeptides, and in certain aspects, isolated novel humanized anti-p185^E^R^G^ antibodies are provided. These novel anti-p185^E^R^G^ antibodies are sometimes collectively referred to herein as huMAb4D5, and also sometimes as the light or heavy chain variable domains of huMAb4D5, and are defined herein to be any polypeptide sequence which possesses a biological property of a polypeptide comprising the following polypeptide sequence:

```
DIQMTQSPSSLSASVGDRVtllTCR4sQD-
```

"Biological property", as relates for example to anti-p185^E^R^G^, for the purposes herein means an in vivo effector function or antigen-binding function or activity that is directly or indirectly performed by huMAb4D5 (whether in its native or denatured conformation) Effector functions include p185^E^R^G^ binding, any hormonal or hormonal antagonist activity, any mitogenic or agonist or antagonist activity, any cytotoxic activity. An antigenic function means possession of an epitope or antigenic site that is capable of cross-reacting with antibodies raised against the polypeptide sequence.

Thus, the biologically active and antigenically active huMAb4D5 polypeptides that are subject of certain embodiments of this invention include the sequence of the entire translated nucleotide sequence of huMAb4D5; mature huMAb4D5; fragments thereof having a consecutive sequence of at least 5, 10, 15, 20, 25, 30 or 40 amino acid residues comprising sequences from muMAb4DS plus residues from the human FR of huMAb4D5; amino acid sequence variants of huMAb4D5 wherein an amino acid residue has been inserted N- or C-terminal to, or within, huMAb4D5 or its fragment as defined above; amino acid sequence variants of huMAb4D5 or its fragment as defined above wherein an amino acid residue of huMAb4D5 or its fragment has been covalently modified, by substitution, chemical, enzymatic, or other appropriate means, with a moiety other than a naturally occurring amino acid; and glycosylation variants of huMAb4D5 (insertion of a glycosylation site or deletion of any glycosylation site by deletion, insertion or substitution of suitable residues). Such fragments and variants exclude any polypeptide hereof identified, including muMAb4DS or any known polypeptides fragment, which are anticipatory order 35 U.S.C. 102 as well as polypeptides obvious thereunder under 35 U.S.C. 103.

An "isolated" polypeptide means polypeptide which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, for example, a polypeptide product comprising huMAb4D5 will be purified from a cell culture or other synthetic environment (1) to greater than 95% by weight of protein as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a gas- or liquid-phase sequenator (such as a commercially available Applied Biosystems sequenator Model 470, 477, or 473), or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Comassie blue or, preferably, silver stain.

Isolated huMAb4D5 includes huMAb4D5 in situ within recombinant cells since at least one component of the huMAb4D5 natural environment will not be present. Ordinarily, however, isolated huMAb4D5 will be prepared by at least one purification step.

In accordance with this invention, huMAb4D5 nucleic acid is RNA or DNA containing greater than ten bases that encodes a biologically or antigenically active huMAb4D5, is complementary to nucleic acid sequence encoding such huMAb4D5, or hybridizes to nucleic acid sequence encoding such huMAb4D5 and remains stably bound to it under stringent conditions, and comprises nucleic acid from a muMAb4DS CDR and a human FR region. Preferably, the huMAb4D5 nucleic acid encodes a polypeptide sharing at least 75% sequence identity, more preferably at least 80%, still more preferably at least 85%, even more preferably at least 90%, and most preferably 95%, with the huMAb4D5 amino acid sequence. Preferably, a nucleic acid molecule that hybridizes to the huMAb4D5 nucleic acid sequence encodes, preferably, a nucleic acid molecule that hybridizes to the huMAb4D5 nucleic acid contains at least 20, preferably 40, and most preferably 90 bases. Such hybridizing or complementary nucleic acid, however, is further defined as being novel under 35 U.S.C. 102 and unobvious under 35 U.S.C. 103 over any prior art nucleic acid.

Stringent conditions are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaC1/0.0015 M sodium citrate/O% NaDodSO4 at 50° C.; (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/O% Ficoll/O% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42° C.; or (3) employ 50% formamide, 5xSSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5xDenhardt’s solution, sonicated salmon sperm DNA (50 g/ml), 0.1% SDS, and 10% dextran sulfate at 42 C., with washes at 42 C. in 0.2xSSC and 0.1% SDS.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, a
A polymerase reaction method for amplifying a nucleic acid test sample, comprising the use of a known nucleic acid (DNA or RNA) as a primer and utilizing a nucleic acid polymerase to amplify or generate a specific piece of nucleic acid or to amplify or generate a specific piece of nucleic acid which is complementary to a particular nucleic acid.

Suitable Methods for Practicing the Invention

Some aspects of this invention include obtaining an import, non-human antibody variable domain, producing a desired humanized antibody sequence and for humanizing an antibody gene sequence are described below. A particularly preferred method of changing a gene sequence, such as gene conversion from a non-human or consensus sequence into a humanized nucleic acid sequence, is the cassette mutagenesis procedure described in Example 1. Additionally, methods are given for obtaining and producing antibodies generally, which apply equally to native non-human antibodies as well as to humanized antibodies.

Generally, the antibodies and antibody variable domains of this invention are conventionally prepared in recombinant cell culture, as described in more detail below. Recombinant synthesis is preferred for reasons of safety and economy, but it is known to prepare peptides by chemical synthesis and to purify them from natural sources; such preparations are included within the definition of antibodies herein.

Molecular Modeling

An integral step in our approach to antibody humanization is construction of computer graphics models of the import and humanized antibodies. These models are used to determine if the six complementarity-determining regions (CDRs) can be successfully transplanted from the import framework to a human one and to determine which framework residues from the import antibody, if any, need to be incorporated into the humanized antibody in order to maintain CDR conformation. In addition, analysis of the sequences of the import and humanized antibodies and reference to the models can help to discern which framework residues are unusual and thereby might be involved in antigen binding or maintenance of proper antibody structure.

All of the humanized antibody models of this invention are based on a single three-dimensional computer graphics structure hereafter referred to as the consensus structure. This consensus structure is a key distinction from the approach of previous workers in the field, who typically begin by selecting a human antibody structure which has an amino acid sequence which is similar to the sequence of their import antibody.

The consensus structure of one embodiment of this invention was built in five steps as described below.

Step 1: Seven Fab X-ray crystal structures from the Brookhaven Protein Data Bank were used (entries 2F84, 2RHE, 3FAB, and 1 REI which are human structures, and 2MCP, 1FBJ, and 2HFL which are murine structures). For each structure, protein mainchain geometry and hydrogen bonding patterns were used to assign each residue to one of three secondary structure types: alpha-helix, beta-strand or other (i.e. non-helix and non-strand). The immunoglobulin residues used in superpositioning and those included in the consensus structure are shown in Table 1.
TABLE I

<table>
<thead>
<tr>
<th></th>
<th>V,K domain</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig'</td>
<td>2FB4</td>
<td>2HHE</td>
<td>2MCP</td>
<td>3FAB</td>
<td>1FBJ</td>
<td>2HPL</td>
<td>1RFJ</td>
</tr>
<tr>
<td>27-37</td>
<td>18-37</td>
<td>18-37</td>
<td>18-37</td>
<td>18-37</td>
<td>18-37</td>
<td>18-37</td>
<td>18-37</td>
</tr>
<tr>
<td>69-72</td>
<td>71-76</td>
<td>76-81</td>
<td>69-74</td>
<td>69-74</td>
<td>70-75</td>
<td>69-74</td>
<td>70-75</td>
</tr>
<tr>
<td>RMSa</td>
<td>0.40</td>
<td>0.60</td>
<td>0.53</td>
<td>0.54</td>
<td>0.48</td>
<td>0.50</td>
<td>0.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>V,H domain</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig'</td>
<td>2FB4</td>
<td>2MCP</td>
<td>3FAB</td>
<td>1FBJ</td>
<td>2HPL</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>46-52</td>
<td>46-52</td>
<td>46-52</td>
<td>46-52</td>
<td>46-52</td>
<td>45-51</td>
<td>3-8</td>
<td></td>
</tr>
<tr>
<td>57-61</td>
<td>59-63</td>
<td>56-60</td>
<td>57-61</td>
<td>57-61</td>
<td>57-61</td>
<td>3-8</td>
<td></td>
</tr>
<tr>
<td>62-71</td>
<td>70-73</td>
<td>67-70</td>
<td>68-71</td>
<td>68-71</td>
<td>66-71</td>
<td>3-8</td>
<td></td>
</tr>
<tr>
<td>76-84</td>
<td>80-86</td>
<td>77-83</td>
<td>78-84</td>
<td>78-84</td>
<td>75-82</td>
<td>3-8</td>
<td></td>
</tr>
<tr>
<td>92-99</td>
<td>94-101</td>
<td>91-98</td>
<td>92-99</td>
<td>95-99</td>
<td>94-101</td>
<td>3-8</td>
<td></td>
</tr>
<tr>
<td>RMSa</td>
<td>0.63</td>
<td>0.85</td>
<td>0.62</td>
<td>0.91</td>
<td>0.33</td>
<td>0.77</td>
<td>0.92</td>
</tr>
</tbody>
</table>

aPDB-letter code for Protein Data Bank file.
bResidue numbers for the crystal structures are taken from the Protein Data Bank files.
Residue numbers for the consensus structure are according to Kabat et al.
Root-mean-square deviation in Å for (N,Ca,C) atoms superimposed on 2FB4
Root-mean-square deviation in Å for (N,Ca,C) atoms superimposed on 2HPL.

Step 2: Having identified the alpha-helices and beta-strands in each of the seven structures, the structures were superimposed on one another using the INSIGHT computer program (Biosym Technologies, San Diego, Calif.) as follows: The 2FB4 structure was arbitrarily chosen as the template (or reference) structure. The 2FB4 was held fixed in space and the other six structures rotated and translated in space so that their common secondary structural elements (i.e., alpha-helices and beta-strands) were oriented such that these common elements were as close in position to one another as possible. (This superpositioning was performed using accepted mathematical formulae rather than actually physically moving the structures by hand.)

Step 3: With the seven structures thus superimposed, for each residue in the template (2FB4) Fab one calculates the distance from the template alpha-carbon atom (Cα) to the analogous Cα atom in each of the other six superimposed structures. This results in a table of Cα—Cα distances for each residue position in the sequence. Such a table is necessary in order to determine which residue positions will be included in the consensus model. Generally, if all Cα—Cα distances for a given residue position were ≤1.0 Å, that position was included in the consensus structure. If for a given position only one Fab crystal structure was >1.0 Å, the position was included but the outlying crystal structure was not included in the next step (for this position only).

In general, the seven β-strands were included in the consensus structure while some of the loops connecting the β-strands, e.g., complementarity-determining regions (CDRs), were not included in view of Cα divergence.

Step 4: For each residue which was included in the consensus structure after step 3, the average of the coordinates for individual mainchain N, Cα, C, O and Cβ atoms were calculated. Due to the averaging procedure, as well as variation in bond length, bond angle and dihedral angle among the crystal structures, this "average" structure contained some bond lengths and angles which deviated from standard geometry. For purposes of this invention, "standard geometry" is understood to include geometries commonly accepted as typical, such as the compilation of bond lengths and angles from small molecule structures in Weiner, S. J. et al., J. Amer. Chem. Soc., 106: 765-784 (1984).

Step 5: In order to correct these deviations, the final step was to subject the "average" structure to 50 cycles of energy minimization (DISCOVER program, Biosym Technologies) using the AMBER (Weiner, S. J. et al., J. Amer. Chem. Soc., 106: 765-784 (1984)) parameter set with only the Cα coordinates fixed (i.e., all other atoms are allowed to move) (energy minimization is described below). This allowed any deviant bond lengths and angles to assume a standard (chemically acceptable) geometry. See Table II.

TABLE II

| Average Bond Lengths and Angles for "Average" (Before) and Energy-Minimized Consensus (After 50 Cycles) Structures |
|---|---|---|---|---|
|  | V,N before (Å) | V,N after (Å) | V,H before (Å) | V,H after (Å) | Standard Geometry (Å) |
| N—Cα | 1.459 | 1.451 | 1.452 | 1.449 |
| Cα—Cα | 1.515 | 1.507 | 1.542 | 1.522 |
| O—Cα | 1.208 | 1.210 | 1.221 | 1.229 |
The consensus structure might conceivably be dependent upon which crystal structure was chosen as the template on which the others were superimposed. As a test, the entire procedure was repeated using the crystal structure with the worst superposition versus 2FB4, i.e. the 2HFL Fab structure, as the new template (reference). The two consensus structures compare favorably (root-mean-squared deviation of 0.11 Å for all N, Ca and C atoms).

Note that the consensus structure only includes mainchain (N, Ca, C, O, CB atoms) coordinates for only those residues which are part of a conformation common to all seven X-ray crystal structures. For the Fab structures, these include the common β-strands (which comprise two β-sheets) and a few non-CDR loops which connect these β-strands. The consensus structure does not include CDRs or sidechains, both of which vary in their conformation among the seven structures. Also, note that the consensus structure includes only the VL and VH domains.

This consensus structure is used as the archetype. It is not particular to any species, and has only the basic shape without side chains. Starting with this consensus structure, the model of any import, human, or humanized Fab can be constructed as follows. Using the amino acid sequence of the particular antibody VL and VH domains of interest, a computer graphics program (e.g. INSIGHT, Biosym Technologies) is used to add sidechains and CDRs to the consensus structure. When a sidechain is added, its conformation is chosen on the basis of known Fab crystal structures (see the Background section for publications of such crystal structures) and rotamer libraries (Fonder, J. W. & Richards, F. M., J. Mol. Biol. 193: 775–791 (1987)). The model also is constructed so that the atoms of the sidechain are positioned so as to not collide with other atoms in the Fab.

CDRs are added to the model (now having the backbone plus side chains) as follows. The size (i.e. number of amino acids) of each import CDR is compared to canonical CDR structures tabulated by Chothia et al., Nature, 342:877–883 (1989) and which were derived from Fab crystals. Each CDR sequence is also reviewed for the presence or absence of certain specific amino acid residues which are identified by Chothia as structurally important: e.g. light chain residues 29 (CDR1) and 55 (CDR3), and heavy chain residues 26, 27, 29 (CDR1) and 55 (CDR2). For light chain CDR2, and heavy chain CDR3, only the size of the CDR is compared to the Chothia canonical structure. If the size and sequence (i.e. inclusion of the specific, structurally important residues as denoted by Chothia et al.) of the import CDR agrees in size and has the same structurally important residues as those of a canonical CDR, then the mainchain conformation of the import CDR in the model is taken to be the same as that of the canonical CDR. This means that the import sequence is assigned the structural configuration of the canonical CDR, which is then incorporated in the evolving model.

However, if no matching canonical CDR can be assigned for the import CDR, then one of two options can be exercised. First, using a program such as INSIGHT (Biosym Technologies), the Brookhaven Protein Data Bank can be searched for loops with a similar size to that of the import CDR and these loops can be evaluated as possible conformations for the import CDR in the model. Minimally, such loops must exhibit a conformation in which no loop atom overlaps with other protein atoms. Second, one can use available programs which calculate possible loop conformations, assuming a given loop size, using methods such as described by Brucoleri et al., Nature 335: 564–568 (1988).

When all CDRs and sidechains have been added to the consensus structure to give the final model (import, human, or humanized), the model is preferably subjected to energy minimization using programs which are available commercially (e.g. DISCOVER, Biosym Technologies). This technique uses complex mathematical formulae to refine the model by performing such tasks as checking that all atoms are within appropriate distances from one another and checking that bond lengths and angles are within chemically acceptable limits.

Models of a humanized, import or human antibody sequence are used in the practice of this invention to understand the impact of selected amino acid residues of the activity of the sequence being modeled. For example, such a model can show residues which may be important in antigen binding, or for maintaining the conformation of the antibody, as discussed in more detail below. Modeling can also be used to explore the potential impact of changing any amino acid residue in the antibody sequence.

Methods for Obtaining a Humanized Antibody Sequence

In the practice of this invention, the first step in humanizing an import antibody is deriving a consensus amino acid sequence into which to incorporate the import sequences. Next a model is generated for these sequences using the methods described above. In certain embodiments of this invention, the consensus human sequences are derived from the most abundant subclasses in the sequence compilation of Kabat et al. (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)), namely Vκ subgroup I and Vκ subgroup III, and have the sequences indicated in the definitions above.

While these steps may be taken in different order, typically a structure for the candidate humanized antibody is created by transferring the at least one CDR from the non-human, import sequence into the consensus human structure, after the entire corresponding human CDR has been removed. The humanized antibody may contain human replacements of the non-human import residues at positions within CDRs as defined by sequence variability (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) or as

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C—N</td>
<td>1.280</td>
<td>1.277(0.002)</td>
<td>1.262</td>
<td>1.275(0.004)</td>
</tr>
<tr>
<td>Cα—Cβ</td>
<td>1.508</td>
<td>1.506(0.002)</td>
<td>1.499</td>
<td>1.508(0.002)</td>
</tr>
</tbody>
</table>

Values in parentheses are standard deviations. Note that while some bond length and angle changes did not change appreciably after energy minimization, the corresponding standard deviations are reduced due to deviation geometries assuming standard values after energy minimization. Standard geometry values are from the AMBER forcefield as implemented in DISCOVER (Biosym Technologies).

**Table II—continued**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C—N—Ca</td>
<td>123.5</td>
<td>123.8(1.1)</td>
<td>123.3</td>
<td>124 (1.1)</td>
</tr>
<tr>
<td>N—Ca—C</td>
<td>116.0</td>
<td>119.5(1.9)</td>
<td>116.3</td>
<td>119.5(1.6)</td>
</tr>
<tr>
<td>Ca—Cα—N</td>
<td>177.1</td>
<td>177.4(0.6)</td>
<td>177.7</td>
<td>177.4(0.4)</td>
</tr>
</tbody>
</table>

While these steps may be taken in different order, typically a structure for the candidate humanized antibody is created by transferring the at least one CDR from the non-human, import sequence into the consensus human structure, after the entire corresponding human CDR has been removed. The humanized antibody may contain human replacements of the non-human import residues at positions within CDRs as defined by sequence variability (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) or as

V_\gamma-CDR1 R24R, V_\gamma-CDR2 R54L, and V_\gamma-CDR3 T65S

Differences between the non-human import and the human consensus framework residues are individually investigated to determine their possible influence on CDR conformation and/or binding to antigen. Investigation of such possible influences is desirably performed through modeling, by examination of the characteristics of the amino acids at particular locations, or determined experimentally through evaluating the effects of substitution or mutagenesis of particular amino acids.

In certain preferred embodiments of this invention, a humanized antibody is made comprising amino acid sequence of an import, non-human antibody and a human antibody, utilizing the steps of:

a. obtaining the amino acid sequences of at least a portion of an import antibody variable domain and of a consensus human variable domain;

b. identifying Complementarity Determining Region (CDR) amino acid sequences in the import and the human variable domain sequences;

c. substituting an import CDR amino acid sequence for the corresponding human CDR amino acid sequence;

d. aligning the amino acid sequences of a Framework Region (FR) of the import antibody and the corresponding FR of the consensus antibody;

e. identifying import antibody FR residues in the aligned FR sequences that are non-homologous to the corresponding consensus antibody residues;

f. determining if the non-homologous import amino acid residue is reasonably expected to have at least one of the following effects:

1. non-covalently binds antigen directly,

2. interacts with a CDR; or

3. participates in the V_\gamma-V_\chi interface; and

g. for any non-homologous import antibody amino acid residue which is reasonably expected to have at least one of these effects, substituting that residue for the corresponding amino acid residue in the consensus antibody FR sequence.

Optionally, one determines if any non-homologous residues identified in step (e) are exposed on the surface of the domain or buried within it, and if the residue is exposed but has none of the effects identified in step (f), one may retain the consensus residue.

Additionally, in certain embodiments the corresponding consensus antibody residues identified in step (e) above are selected from the group consisting of 4L, 35L, 38L, 36L, 43L, 4GL, 58L, 4GL, 62L, 63L, 64L, 66L, 69L, 67L, 70L, 71L, 73L, 75L, 76L, 78L, 91L, 92L, 93L, and 103L.

Another preferred embodiment of the methods of this invention comprises aligning import antibody and the consensus antibody FR sequences, identifying import antibody FR residues which are non-homologous with the aligned consensus FR sequence, and for each such non-homologous import antibody FR residue, determining if the corresponding consensus antibody residue represents a residue which is highly conserved across all species at that site, and if it is so conserved, preparing a humanized antibody which comprises the consensus antibody amino acid residue at that site.

In certain alternate embodiments, one need not utilize the modeling and evaluation steps described above, and may instead proceed with the steps of obtaining the amino acid sequence of at least a portion of an import, non-human antibody variable domain having a CDR and FR, obtaining the amino acid sequence of at least a portion of a consensus human antibody variable domain having a CDR and FR, substituting the non-human CDR for the human CDR in the consensus human antibody variable domain, and then substituting an amino acid residue for the consensus amino acid residue at at least one of the following sites:


Preferably, the non-CDR residue substituted at the consensus FR site is the residue found at the corresponding location of the non-human antibody. If desired, one may utilize the other method steps described above for determining whether a particular amino acid residue can reasonably be expected to have undesirable effects, and remedying those effects.

If after making a humanized antibody according to the steps above and testing its activity one is not satisfied with the humanized antibody, one preferably reexamines the potential effects of the amino acids at the specific locations recited above. Additionally, it is desirable to reinvestigate any buried residues which are reasonably expected to affect the V_\gamma-V_\chi interface but may not directly affect CDR conformation. It is also desirable to reevaluate the humanized antibody utilizing the steps of the methods claimed herein.

In certain embodiments of this invention, amino acid residues in the consensus human sequence are substituted for by other amino acid residues. In preferred embodiments, residues from a particular non-human import sequence are substituted, however there are circumstances where it is desired to evaluate the effects of other amino acids. For example, if after making a humanized antibody according to the steps above and testing its activity one is not satisfied...
with the humanized antibody, one may compare the sequences of other classes or subgroups of human antibodies, or classes or subgroups of antibodies from the particular non-human species, and determine which other amino acid side chains and amino acid residues are found at particular locations and substituting such other residues.

Antibodies

Certain aspects of this invention are directed to natural antibodies and to monoclonal antibodies, as illustrated in the Examples below and by antibody hybridomas deposited with the ATCC (as described below.) Thus, the references throughout this description to the use of monoclonal antibodies are intended to include the use of natural or native antibodies as well as humanized and chimeric antibodies. As used herein, the term “antibody” includes the antibody variable domain and other separable antibody domains unless specifically excluded.

In accordance with certain aspects of this invention, antibodies to be humanized (import antibodies) are isolated from continuous hybrid cell lines formed by the fusion of antigen-primed immune lymphocytes with myeloma cells.

In certain embodiments, the antibodies of this invention are obtained by routine screening. Polyclonal antibodies to an antigen generally are isolated from animals by multiple subcutaneous (sc) or intraperitoneal (ip) injections of the antigen and an adjuvant. It may be useful to conjugate the antigen or a fragment containing the target antigenic sequence to a protein that is immunogenic in the species to be immunized, e.g., keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, or soybean trypsin inhibitor using a bifunctional or derivatizing agent, for example, maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride, SOCl₂, or R'-N=C=NR, where R and R' are different alkyl groups.

The route and schedule of the host animal or cultured antibody-producing cells therefrom are generally in keeping with established and conventional techniques for antibody stimulation and production. While mice are frequently employed as the test model, it is contemplated that any mammalian subject including human subjects or antibody-producing cells obtained therefrom can be manipulated according to the processes of this invention to serve as the basis for production of mammalian, including human, hybrid cell lines.

Animals are typically immunized against the immunogenic conjugates or derivatives by combining 1 mg or 1 µg of conjugate (for rabbits or mice, respectively) with 3 volumes of Freund’s complete adjuvant and injecting the solution intradermally at multiple sites. One month later the animals are boosted with ¼ to ½ of the original amount of conjugate in Freund’s complete adjuvant (or other suitable adjuvant) by subcutaneous injection at multiple sites. 7 to 14 days later animals are bled and the sera is assayed for antigen titer. Animals are boosted until the titer plateaus. Preferably, the animal is boosted with the conjugate of the same antigen, but conjugated to a different protein and/or through a different crosslinking agent. Conjugates also can be made in recombinant cell culture as protein fusions. Also, aggregating agents such as alum are used to enhance the immune response.

After immunization, monoclonal antibodies are prepared by recovering immune lymphoid cells—typically spleen cells or lymphocytes from lymph node tissue—from immunized animals and immortalizing the cells in conventional fashion, e.g., by fusion with myeloma cells or by Epstein-Barr (EB) virus transformation and screening for clones expressing the desired antibody. The hybridoma technique described originally by Kohler and Milstein, Eur. J. Immunol. 6:511 (1976) has been widely applied to produce hybrid cell lines that secrete high levels of monoclonal antibodies against many specific antigens.

It is possible to fuse cells of one species with another. However, it is preferable that the source of the immunized antibody producing cells and the myeloma be from the same species.

The hybrid cell lines can be maintained in culture in vitro in cell culture media. The cell lines of this invention can be selected and/or maintained in a composition comprising the continuous cell line in hypoxanthine-aminopterine-thymidine (HAT) medium. In fact, once the hybridoma cell line is established, it can be maintained on a variety of nutritionally adequate media. Moreover, the hybrid cell lines can be stored and preserved in any number of conventionally way, including freezing and storage under liquid nitrogen. Frozen cell lines can be revived and cultured indefinitely with resumed synthesis and secretion of monoclonal antibody.

The secreted antibody is recovered from tissue culture supernatant by conventional methods such as precipitation, ion exchange chromatography, affinity chromatography, or the like. The antibody described herein are also recovered from lyophilized cell cultures by conventional methods for purification of IgG or IgM as the case may be that hereofore have been used to purify these immunoglobulins from pooled plasma, e.g. ethanol or polyethylene glycol precipitation procedures. The purified antibodies are sterile filtered, and optionally are conjugated to a detectable marker such as an enzyme or spin label for use in diagnostic assays of the antigen in test samples.

While routinely used monoclonal antibodies are used as the source of the import antibody, the invention is not limited to any species. Additionally, techniques developed for the production of chimeric antibodies (Morrison et al., Proc. Natl. Acad. Sci., 81:6851 (1984); Neuberger et al., Nature 312:604 (1984); Takada et al., Nature 314:452 (1985)) by splicing the genes from a vaccine antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity (such as ability to activate human complement and mediate ADCC) can be used; such antibodies are within the scope of this invention.

Techniques for creating recombinant DNA versions of the antigen-binding regions of antibody molecules (known as Fab fragments) which bypass the generation of monoclonal antibodies are encompassed within the practice of this invention. One extracts antibody-specific messenger RNA molecules from immune system cells taken from an immunized animal, transcribes these into complementary DNA (cDNA), and clones the cDNA into a bacterial expression system. One example of such a technique suitable for the practice of this invention was developed by researchers at Scripps/Stratagene, and incorporates a proprietary bacteriophage lambda vector system which contains a leader sequence that causes the expressed Fab protein to migrate to the periplasmic space (between the bacterial cell membrane and the cell wall) or to be secreted. One can rapidly generate and screen great numbers of functional Fab fragments for those which bind the antigen. Each Fab fragment's specificity for the antigen are specifically encompassed within the term “antibody” as it is defined, discussed, and claimed herein.

Amino Acid Sequence Variants

Amino acid sequence variants of the antibodies and polypeptides of this invention (referred to herein as the
target polypeptide are prepared by introducing appropriate nucleotide changes into the DNA encoding the target polypeptide, or by in vitro synthesis of the desired target polypeptide. Such variants include, for example, humanized variants of non-human antibodies, as well as deletions from, or insertions of, or substitutions of residues within particular amino acid sequences. Any combination of deletion, insertion, and substitution can be made to arrive at the final construct, provided that the final construct possesses the desired characteristics. The amino acid changes may alter or introduce, or substitute for, the sites of substitution, for example, as a sheet or helix. Naturally occurring residues are divided into groups based on common side chain properties: (1) hydrophobic: norleucine, met, ala, val, leu, ile; (2) neutral hydrophilic: cys, ser, thr; (3) acidic: asp, glu; (4) basic: arg, lirn, lys, arg; (5) residues that influence chain orientation: g1y, pro, and (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a neutral hydrophilic residue for a charged amino acid residue, an artifact of the direct expression of target polypeptide in bacterial recombinant cell culture, and fusion of a heterologous N-terminal signal sequence to the N-terminus of the target polypeptide molecule to facilitate the secretion of the mature target polypeptide from recombinant host cells. Such signal sequences generally will be obtained from, and thus homologous to, the intended host cell species. Suitable signal sequences include StH or lpp for E. coli, alpha factor for yeast, and viral signals, such as herpes gD for mammalian cells.

Other insertion variants of the target polypeptide include the fusion to the N- or C-terminus of the target polypeptide of immunogenic polypeptides, e.g., bacterial polysaccharides such as beta-lactamase or an enzyme encoded by the E. coli trp loci, or yeast protein, and C-terminal fusions with proteins having a long half-life such as immunoglobulin constant regions (or other immunoglobulin regions), albumin, or ferritin, as described in WO 89/02922 published Apr. 6, 1989.

Another group of variants are amino acid substitution variants. These variants have at least one amino acid residue in the target polypeptide molecule removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include sites identified as the active site(s) of the target polypeptide, and sites where the amino acids found in the target polypeptide from various species are substantially different in terms of side-chain bulk, charge, and hydrophobicity. Other sites for substitution are described infra, considering the effect of the substitution of the antigen binding, affinity and other characteristics of a particular target antibody.

Other sites of interest are those in which particular residues of the target polypeptides observed from various species are identical. These positions may be important for the biological activity of the target polypeptide. These sites, especially those falling within a sequence of at least three other identical conserved sites, are substituted in a relatively conservative manner. If such substitutions result in a change in biological activity, then other changes are introduced and the products screened until the desired effect is obtained.

Substantial modifications in function or immunological identity of the target polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side chain properties: (1) hydrophobic: norleucine, met, ala, val, leu, ile; (2) neutral hydrophilic: cys, ser, thr; (3) acidic: asp, glu; (4) basic: arg, lirn, lys, arg; (5) residues that influence chain orientation: g1y, pro; and (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will catalyze exchanging a member of one of these classes for another. Such substituted residues may be introduced into regions of the target polypeptide that are homologous with other antibodies of the same class or subclass, or, more preferably, into the non-homologous regions of the molecule.

Any cysteine residues not involved in maintaining the proper conformation of target polypeptide may be substituted, genetically engineered, to improve the oxidative stability of the molecule and prevent aberrant cross-linking.

DNA encoding amino acid sequence variants of the target polypeptide is prepared by a variety of methods known in
the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of the target polypeptide. A particularly preferred method of gene conversion mutagenesis is described below in Example 1. These techniques may utilize target polypeptide nucleic acid (DNA or RNA), or nucleic acid complementary to the target polypeptide nucleic acid.

Oligonucleotide-mediated mutagenesis is a preferred method for preparing substitution, deletion, and insertion variants of target polypeptide DNA. This technique is well known in the art as described by Adelman et al., DNA, 2: 183 (1983). Briefly, the target polypeptide DNA is altered by hybridizing an oligonucleotide encoding the desired mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of the target polypeptide. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the target polypeptide DNA.

Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea et al. (Proc. Natl. Acad. Sci. USA, 75: 5765 [1978]).

Single-stranded DNA template may also be generated by denaturing double-stranded plasmid (or other) DNA using standard techniques. For alteration of the native DNA sequence (to generate amino acid sequence variants, for example), the oligonucleotide is hybridized to the single-stranded template under suitable hybridization conditions. A DNA polymerizing enzyme, usually the Klenow fragment of DNA polymerase 1, is then added to synthesize the complementary strand of the template using the oligonucleotide as a primer for synthesis. A heteroduplex molecule is thus formed such that one strand of DNA encodes the mutated form of the target polypeptide, and the other strand (the original template) encodes the native, unaltered sequence of the target polypeptide. This heteroduplex molecule is then transformed into a suitable host cell, usually a prokaryote such as E. coli JM101. After the cells are grown, they are plated onto agarose plates and screened using the oligonucleotide primer radiolabeled with 32-phosphate to identify the bacterial colonies that contain the mutated DNA. The mutated region is then removed and placed in an appropriate vector for transformation of an appropriate host.

The method described immediately above may be modified such that a homoduplex molecule is created wherein both strands of the plasmid contain the mutation(s). The modifications are as follows: The single-stranded oligonucleotide is annealed to the single-stranded template as described above. A mixture of three deoxyribonucleotides, deoxyribodideoxynucleosine triphosphates, ATP, and DNA ligase. This homoduplex molecule can then be transformed into a suitable host cell such as E. coli JM101, as described above.

DNA encoding target polypeptide variants with more than one amino acid to be substituted may be generated in one of several ways. If the amino acids are located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If, however, the amino acids are located some distance from each other (separated by more than about ten amino acids), it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed.

In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded template DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions.

The alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: wild-type DNA is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more mutations. The oligonucleotide encoding the additional desired amino acid substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. This resultant DNA can be used as a template in a third round of mutagenesis, and so on.

PCR mutagenesis is also suitable for making amino acid variants of target polypeptide. While the following discussion refers to DNA, it is understood that the technique also finds application with RNA. The PCR technique generally refers to the following procedure (see Erlich, supra, the chapter by R. Higuchi, p. 61-70). When small amounts of template DNA are used as starting material in a PCR, primers that differ slightly in sequence from the corresponding region in a template DNA can be used to generate relatively large quantities of a specific DNA fragment that differs from the template sequence only at the positions where the primers differ from the template. For introduction of a mutation into a plasmid DNA, one of the primers is designed to overlap the position of the mutation and to contain the mutation; the sequence of the other primer must be identical to a stretch of sequence of the opposite strand of the plasmid, but this sequence can be obtained from Amersham Corporation. This mixture is added to the template-oligonucleotide complex. Upon addition of DNA polymerase to this mixture, a strand of DNA identical to the template except for the mutated bases is generated. In addition, this new strand of DNA will contain dCTP-(as) instead of dCTP, which serves to protect it from restriction endonuclease digestion.

After the template strand of the double-stranded heteroduplex is nicked with an appropriate restriction enzyme, the template strand can be digested with ExoIII nuclease or another appropriate nuclease past the region that encircles the site(s) to be mutagenized. The reaction is then stopped to leave a molecule that is only partially single-stranded. A complete double-stranded DNA homoduplex is then formed using DNA polymerase in the presence of all four deoxyribonucleotide triphosphates, ATP, and DNA ligase. This homoduplex molecule can then be transformed into a suitable host cell such as E. coli JM101, as described above.

DNA encoding target polypeptide variants with more than one amino acid to be substituted may be generated in one of several ways. If the amino acids are located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If, however, the amino acids are located some distance from each other (separated by more than about ten amino acids), it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed.

In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded template DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions.

The alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: wild-type DNA is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more mutations. The oligonucleotide encoding the additional desired amino acid substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. This resultant DNA can be used as a template in a third round of mutagenesis, and so on.

PCR mutagenesis is also suitable for making amino acid variants of target polypeptide. While the following discussion refers to DNA, it is understood that the technique also finds application with RNA. The PCR technique generally refers to the following procedure (see Erlich, supra, the chapter by R. Higuchi, p. 61-70). When small amounts of template DNA are used as starting material in a PCR, primers that differ slightly in sequence from the corresponding region in a template DNA can be used to generate relatively large quantities of a specific DNA fragment that differs from the template sequence only at the positions where the primers differ from the template. For introduction of a mutation into a plasmid DNA, one of the primers is designed to overlap the position of the mutation and to contain the mutation; the sequence of the other primer must be identical to a stretch of sequence of the opposite strand of the plasmid, but this sequence can be located anywhere.
along the plasmid DNA.

It is preferred, however, that the sequence of the second primer is located within 200 nucleotides from that of the first, such that in the end the entire amplified region of DNA bounded by the primers can be easily sequenced. PCR amplification using a primer pair like the one just described results in a population of DNA fragments that differ at the position of the mutation specified by the primer, and possibly at other positions, as template copying is somewhat error-prone.

If the ratio of template to product material is extremely low, the vast majority of product DNA fragments incorporate the desired mutation(s). This product material is used to replace the corresponding region in the plasmid that served as PCR template using standard DNA technology. Mutations at separate positions can be introduced simultaneously by either using a mutant second primer, or performing a second PCR with different mutant primers and ligating the two resulting PCR fragments simultaneously to the vector fragment in a three (or more)-part ligation.

In a specific example of PCR mutagenesis, template plasmid DNA (1 µg) is linearized by digestion with a restriction endonuclease that has a unique recognition site in the plasmid DNA outside of the region to be amplified. Of this material, 100 ng is added to a PCR mixture containing PCR buffer, which contains the four deoxynucleotide triphosphates and is included in the GeneAmp® kits (obtained from Perkin-Elmer Cetus, Norwalk, Conn. and Emeryville, Calif.), and 25 pmole of each oligonucleotide primer, to a final volume of 50 µl. The reaction mixture is overlayed with 35 µl mineral oil. The reaction is denatured for 5 minutes at 100°C, placed briefly on ice, and then 1 µl Thermus aquaticus (Taq) DNA polymerase (5 units/µl, purchased from Perkin-Elmer Cetus, Norwalk, CT and Emeryville, Calif.) is added below the mineral oil layer. The reaction mixture is then inserted into a DNA Thermal Cycler (purchased from Perkin-Elmer Cetus) programmed as follows: 2 min. at 95°C, then 30 sec. at 72°C, then 19 cycles of the following: 30 sec. at 94°C, 30 sec. at 55°C, and 30 sec. at 72°C.

At the end of the program, the reaction vial is removed from the thermal cycler and the aqueous phase transferred to a new vial, extracted with phenol/chloroform (50:50 vol.), and ethanol precipitated, and the DNA is recovered by standard procedures. This material is subsequently subjected to the appropriate treatments for insertion into a vector.

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells et al. (Gene, 34: 515 [1983]). The starting material is the plasmid (or other vector) comprising the target polypeptide DNA to be mutated. The codon(s) in the target polypeptide DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the target polypeptide DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is referred to as the cassette. This cassette is designed to have 3' and 5' ends that are compatible with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated target polypeptide DNA sequence.

Insertion of DNA into a Cloning Vehicle

The cDNA or genomic DNA encoding the target polypeptide is inserted into a replicable vector for further cloning (amplification of the DNA) or for expression. Many vectors are available, and selection of the appropriate vector will depend on 1) whether it is to be used for DNA amplification or for DNA expression, 2) the size of the DNA to be inserted into the vector, and 3) the host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of DNA or expression of DNA) and the host cell for which it is compatible.

The vector components generally include, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence.

(a) Signal Sequence Component

In general, the signal sequence may be a component of the vector, or it may be a part of the target polypeptide DNA that is inserted into the vector.

The target polypeptides of this invention may be expressed not only directly, but also as a fusion with a heterologous polypeptide, preferably a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the target polypeptide DNA that is inserted into the vector. Included within the scope of this invention are target polypeptides with any native signal sequence deleted and replaced with a heterologous signal sequence.

The heterologous signal sequence selected should be one that is recognized and processed (i.e. cleaved by a signal peptidase) by the host cell. For prokaryotic host cells that do not recognize and process the native target polypeptide signal sequence, the signal sequence is substituted by a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, βp, or heat-stable enterotoxin II leaders. For yeast secretion the native target polypeptide signal sequence may be substituted by the yeast invertase, α-factor, or acid phosphatase leaders. In mammalian cell expression the native signal sequence is satisfactory, although other mammalian signal sequences may be suitable.

(b) Origin of Replication Component

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Generally, in cloning vectors this sequence is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (the SV40 origin may typically be used only because it contains the early promoter).

Most expression vectors are “shuttle” vectors, i.e. they are capable of replication in at least one class of organisms but can be transferred into another organism for expression. For example, a vector is a plasmid in E. coli and then the same vector is transfected into yeast or mammalian cells for expression even though it is not capable of replicating independently of the host cell chromosome.

DNA may also be amplified by insertion into the host genome. This is readily accomplished using Bacillus species
as hosts, for example, by including in the vector a DNA sequence that is complementary to a sequence found in Bacillus genomic DNA. Transfection of Bacillus with this vector results in homologous recombination with the genome and insertion of the target polypeptide DNA. However, the recovery of genomic DNA encoding the target polypeptide is more complex than that of an exogenously replicated vector because restriction enzyme digestion is required to excise the target polypeptide DNA.

(c) Selection Gene Component

Expression and cloning vectors should contain a selection gene, also termed a selectable marker. This gene encodes a protein necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxic agents, e.g. ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g. the gene encoding D-alanine racemase for Bacillus.

One example of a selection scheme utilizes a drug to arrest growth of a host cell. Those cells that are successfully transformed with a heterologous gene express a protein conferring drug resistance and thus survive the selection regimen. Examples of such dominant selection use the drugs neomycin (Southern et al., J. Mol. Appl. Genet., 1: 327 [1982]), mycophenolic acid (Mulligan et al., Science, 209: 1422 [1980]) or hygromycin (Boguski et al., Mol. Cell. Biol., 5: 410-413 [1985]). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid), or hygromycin, respectively.

Another example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the target polypeptide nucleic acid, such as dihydrofolate reductase (DHFR) or thymidine kinase. The mammalian cell line used for selection pressure which only the transformants are uniquely adapted to survive by virtue of having taken up the marker. Selection pressure is imposed by culturing the transformants under conditions in which the concentration of selection agent in the medium is successively changed, thereby leading to amplification of both the selection gene and the DNA that encodes the target polypeptide. Amplification of the process by which genes in greater demand for the production of a protein critical for growth are reiterated in tandem within the chromosomes of successive generations of recombinant cells. Increased quantities of the target polypeptide are synthesized from the amplified DNA.

For example, cells transformed with the DHFR selection gene are first identified by culturing all of the transformants in a culture medium that contains methotrexate (Mtx), a competitive antagonist of DHFR. An appropriate host cell when wild-type DHFR is employed is the Chinese hamster ovary (CHO) cell line deficient in DHFR activity, prepared and propagated as described by Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77: 4216 [1980]. The transformed cells are then exposed to increased levels of methotrexate. This leads to the synthesis of multiple copies of the DHFR gene and, concomitantly, multiple copies of other DNA comprising the expression vectors, such as the DNA encoding the target polypeptide. This amplification technique can be used with any otherwise suitable host, e.g., ATCC No. CCL61 CHO-K1, notwithstanding the presence of endog-
Suitable promoter sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem., 255: 2073 [1980]) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg., 7: 149 [1968]; and Holland, Biochemistry, 17: 4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucone isomerase, and glutokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in Hitzeman et al., ET, 269 (1980). Yeast enhancers also are advantageous when used with yeast promoters.

Promoter sequences are known for eukaryotes. Virtually all eukaryotic genes have an AI-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXXAAT region where X may be any nucleotide. At the 3' end of most eukaryotic genes is an AGATAAA sequence that may be the signal for addition of the poly A tail to the 3' end of the coding sequence. All of these sequences are suitably inserted into mammalian expression vectors.

Target polypeptide transcription from vectors in mammalian host cells is controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published Jul. 5, 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g. the actin promoter or an immunglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the target polypeptide sequence, provided such promoters are compatible with the host cell systems.

The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication. Fiers et al., Nature, 273:113 (1978); Mulligan and Berg, Science, 209: 1422-1427 (1980); Pavlakis et al., Proc. Natl. Acad. Sci. USA, 78: 993 [1981] and 3' (Lusky et al., Mol. Cell Bio., 3: 1108 [1983]) to the transcription unit, within an intron (Banerji et al., Cell, 33: 129 [1983]) as well as within the coding sequence itself (Oshorn et al., Mol. Cell Bio., 4:1239 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. See also Yagi, Nature, 297: 17-18 (1982) on enhancing elements for activation of eukaryotic promoters. The enhancer may be spliced into the vector at a position 5' or 3' to the target polypeptide DNA, but is preferably located at a site 5' from the promoter.

Promoter sequences are known for eukaryotes. Virtually all eukaryotic genes have an AI-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXXAAT region where X may be any nucleotide. At the 3' end of most eukaryotic genes is an AGATAAA sequence that may be the signal for addition of the poly A tail to the 3' end of the coding sequence. All of these sequences are suitably inserted into mammalian expression vectors.

Target polypeptide transcription from vectors in mammalian host cells is controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published Jul. 5, 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g. the actin promoter or an immunglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the target polypeptide sequence, provided such promoters are compatible with the host cell systems.

The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication. Fiers et al., Nature, 273:113 (1978); Mulligan and Berg, Science, 209: 1422-1427 (1980); Pavlakis et al., Proc. Natl. Acad. Sci. USA, 78: 993 [1981] and 3' (Lusky et al., Mol. Cell Bio., 3: 1108 [1983]) to the transcription unit, within an intron (Banerji et al., Cell, 33: 129 [1983]) as well as within the coding sequence itself (Oshorn et al., Mol. Cell Bio., 4:1239 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. See also Yagi, Nature, 297: 17-18 (1982) on enhancing elements for activation of eukaryotic promoters. The enhancer may be spliced into the vector at a position 5' or 3' to the target polypeptide DNA, but is preferably located at a site 5' from the promoter.

Construction of suitable vectors containing one or more of the above listed components the desired coding and control sequences employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and recombined in the form desired to generate the plasmids required.

For analysis to confirm correct sequences in plasmids constructed, the ligation mixtures are used to transform E. coli K12 strain 294 (ATCC 31,446) and successful transformants selected by ampicillin or tetracycline resistance where appropriate. Plasmids from the transformants are prepared, analyzed by restriction endonuclease digestion, and/or sequenced by the method of Messing et al., Nucleic Acids Res., 9: 1891 (1981) or by the method of Maxam et al., Methods in Enzymology, 65: 499 (1980).

Particularly useful in the practice of this invention are expression vectors that provide for the transient expression in mammalian cells of DNA encoding the target polypeptide. In general, transient expression involves the use of an expression vector that is able to replicate efficiently in a host cell, such that the host cell accumulates many copies of the expression vector and, in turn, synthesizes high levels of a desired polypeptide encoded by the expression vector. Transient expression systems, comprising a suitable expression vector and a host cell, allow for the convenient isolation of the target polypeptide in the form desired and for the study of the effects of the polypeptide in vivo upon the cell. Some of the many expression vectors and host cells that have been shown to be useful in the practice of this invention include: (a) plasmid vectors, such as plasmid pBR322, which are replicating in host cell cultures; (b) yeast vectors, which are replicating in yeast strains; (c) baculovirus vectors, (d) liposomes, and (e) viruses such as adenovirus, simian virus 40, herpes simplex virus, vesicular stomatitis virus, and bovine papilloma virus. The expression vectors are transfected into the target host cell line using any suitable recombinant techniques.

Selection and Transformation of Host Cells

Suitable host cells for cloning or expressing the vectors herein are the prokaryotic, yeast, or higher eukaryotic cells described above. Suitable prokaryotes include eubacteria, such as Gram-negative or Gram-positive organisms, for example, E. coli, Bacilli such as S. subtilis, Pseudomonas species such as P. aeruginosa, Salmonella typhimurium, or Serratia marcescens. One preferred E. coli cloning host is E. coli 294 (ATCC 31,446), although other strains such as E. coli B, E. coli 2177 (ATCC 31,537), and E. coli W3110 (ATCC 27,325) are suitable. These examples are illustrative rather than limiting. Preferably the host cell should secrete minimal amounts of proteolytic enzymes. Alternatively, in vitro methods of cloning, e.g. PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable hosts for target polypeptide-encoding vectors. Saccharomyces cerevisiae, or common baker's yeast, is one of the most commonly used among lower eukaryotic host microorganisms. However, a number of other genera, species, and strains are commonly available and useful herein, such as Schizosaccharomyces pombe [Beach and Nurse, Nature, 290: 140 (1981); EP 139,383 published May 2, 1985], Kluyveromyces lactis (U.S. Pat. No. 4,943,529) such as, e.g., K. lactis [Louvencourt et al., J. Bacteriol., 171 (1983)]; K. fragilis, K. bulgaricus, K. thermobifluorans, and K. thermotolerans [Fichter et al., J Basic Microbiol., 28: 265–278 (1988)], Pichia pastoris [EP 183,070; Sreekrishna et al., J. Basic Microbiol., 18: 255–278 (1988)], C. trichoderma reesiae [EP 244,234], Neurospora crassa [Case et al., Proc. Natl. Acad. Sci. USA, 76: 5259–5263 (1979)], and filamentous fungi such as, e.g., Neurospora, Penicillium, Trichoderma, and filamentous fungi such as, e.g., Aspergillus niger [Kelly and Hynes, EMBO J. 4: 475–479 (1985)].

Suitable host cells for the expression of glycosylated target polypeptide are derived from multicellular organisms. Such host cells are capable of complex processing and glycosylation activities. In principle, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture. Examples of invertebrate cells include plant and insect cells. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as Spodoptera frugiperda (caterpillar), Aedes aegypti (mosquito), Ades aegypti (mosquito), Drosophila melanogaster (fruitfly), and Bombyx mori host cells have been identified. See, e.g., Juckow et al., Bio/Technology 6: 47-55 (1988); Miller et al., in Genetic Engineering, Setlow, J. K. et al., eds., Vol. 8 (Plenum Publishing, 1986), pp. 277-279; and Masse et al., Nature, 315: 592–594 (1985). A variety of such viral strains are publicly available, e.g., the L-1 variant of Autographa californica NPV and the Bm-5 strain of Bombyx mori NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of Spodoptera frugiperda cells. Plant cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco can be utilized as hosts. Typically, plant cells are transfected by incubation with certain strains of the baculovirus Agrobacterium tumefaciens, which has been previously manipulated to contain the target polypeptide DNA. During incubation of the plant cell culture with A. tumefaciens, the DNA encoding target polypeptide is transferred to the plant cell host such that it is transfected, and will, under appropriate conditions, express the target polypeptide DNA. In addition, regulatory and signal sequences compatible with plant cells are available, such as the nopaline synthase promoter and polyadenylation signal sequences. Depicker et al., J. Mol. Appl. Gen., 1: 561 (1982).

In addition, DNA sequences isolated from the upstream region of the T-DNA 780 gene are capable of activating or increasing transcription levels of plant-expressible genes in recombinant DNA-containing plant tissues. See EP 321,196 published Jun. 21, 1989.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in tissue culture has become a routine procedure in recent years [Tissue Culture, Academic Press, Knuse and Patterson, editors (1973)]. Examples of useful mammalian host cell lines are monkey kidney CV line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line 293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen. Virol., 36: 59 [1977]; baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells (CHO, Urbach and Quilin, Proc. Natl. Acad. Sci. USA, 77: 4216 [1980]); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23: 243–251 [1980]); monkey kidney cells (CVI ATCC CCL 70); African green monkey kidney cells (MRC-2, ATCC CRL-1378), human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CCL 1442); human lung cells (WI-38, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL 51); TRI cells (Mather et al., Annals N.Y. Acad. Sci., 383: 44–68 [1982]); MRC 5 cells; FS4 cells; and a human hepatoma cell line (Hep G2).

Preferred host cells are human embryonic kidney 293 and Chinese hamster ovary cells.

Host cells are transfected and preferably transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Transfection refers to the taking up of an expression vector by a host cell whether or not any coding sequences are in fact expressed. Numerous methods of transfection are known to the ordinarily skilled artisan, for example, CaPO4 and electroporation. Successful transfection is generally recognized when any indication of the operation of this vector occurs within the host cell.

Transformation means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integrant. Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment technique is a routine procedure in recent years. Transformation of plant cells using Agrobacterium tumefaciens is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23: 315 (1983) and WO 89/05859 published Jun. 29, 1980. For mammalian cells without such cell walls, the calcium phosphate precipitation method described in sections 16.30–16.37 of Sambrook et al., supra, is preferred. General aspects of mammalian cell host system transforma-
Acad. Sci. (USA), 76: 3829 (1979). However, other methods for introducing DNA into cells such as by nuclear injection, electroporation, or protoplast fusion may also be used.

Culturing the Host Cells

Prokaryotic cells used to produce the target polypeptide of this invention are cultured in suitable media as described generally in Sambrook et al., supra.

The mammalian host cells used to produce the target polypeptide of this invention may be cultured in a variety of media. Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium ([MEM], Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ([MEM], Sigma) are suitable for culturing the host cells. In addition, any of the media described in Ham and Wallace, Meth. Enz., 58: 44 (1979), Barnes and Sato, Anal. Biochem., 102: 255 (1980), U.S. Pat. Nos. 4,767,704; 4,927,762; 5,033,582; and 5,033,583, and as commercially available from Sigma, Gentamicin™ drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The host cells referred to in this disclosure encompass cells in vitro culture as well as cells that are within a host animal.

It is further envisioned that the target polypeptides of this invention may be produced by homologous recombination, or with recombinant production methods utilizing ceont elements introduced into cells already containing DNA encoding the target polypeptide currently in use in the field. For example, a powerful promoter/enhancer element, a suppressor, or an exogenous transcripion modulatory element is inserted in the genome of the intended host cell in proximity and orientation sufficient to influence the transcription of DNA encoding the desired target polypeptide. The ceont elenent does not encode the target polypeptide of this invention, but the DNA is present in the host cell genome. One next screen for cells making the target polypeptide of this invention, or increased or decreased levels of expression, as desired.

Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77: 5201-5205 [1980]), dot blotting (DNA analysis), or in situ hybridization, using an appropriately labeled probe, based on the sequences provided herein. Various labels may be employed, most commonly radioisotopes, particularly 35S. However, other techniques may also be employed, such as using biotin-modified nucleotides for introduction into a poly nucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorosceners, enzymes, or the like. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunologichal methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product.

With immunohistochemical staining techniques, a cell sample is prepared, typically by dehydration and fixation, followed by reaction with labeled antibodies specific for the gene product coupled, where the labels are usually visually detectable, such as enzymatic labels, fluorescent labels, luminescent labels, and the like. A particularly sensitive staining technique suitable for use in the present invention is described by Hsu et al., Am. J. Clin. Path., 75: 734-738 (1980).

Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native target polypeptide or against a synthetic peptide based on the DNA sequences provided herein as described further in Section 4 below.

Purification of the Target Polypeptide

The target polypeptide preferably is recovered from the culture medium as a secreted polypeptide, although it also may be recovered from host cell lysates when directly expressed without a secretory signal.

When the target polypeptide is expressed in a recombinant cell other than one of human origin, the target polypeptide is completely free of proteins or polypeptides of human origin. However, it is necessary to purify the target polypeptide from recombinant cell proteins or polypeptides to obtain preparations that are substantially homogeneous to the target polypeptide. As a first step, the culture medium or lysate is centrifuged to remove particulate cell debris. The membrane and soluble protein fractions are then separated. The target polypeptide may then be purified from the soluble protein fraction and from the membrane fraction of the culture lysate, depending on whether the target polypeptide is membrane bound. The following procedures are exemplary of suitable purification procedures: fractionation on immunoadfinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; and protein A Sepharose columns to remove contaminants such as IgG.

Target polypeptide variants in which residues have been deleted, inserted or substituted are recovered in the same fashion, taking account of any substantial changes in properties occasioned by the variation. For example, preparation of a target polypeptide fusion with another protein or polypeptide, e.g. a bacterial or viral antigen, facilitates purification; an immunoaffinity column containing antibody to the antigen (or containing antigen, where the target polypeptide is an antibody) can be used to adsorb the fusion. Immunoaffinity columns such as a rabbit polyclonal anti-target polypeptide column can be employed to absorb the target polypeptide variant by binding it to at least one
remaining immune epitope. A protease inhibitor such as phenyl methyl sulfonyl fluoride (PMSF) also may be useful to inhibit proteolytic degradation during purification, and antibiotics may be included to prevent the growth of adventitious contaminants. One skilled in the art will appreciate that purification methods suitable for native target polypeptide may require modification to account for changes in the character of the target polypeptide or its variants upon expression in recombinant cell culture.

Covalent Modifications of Target Polypeptides

Covalent modifications of target polypeptides are included within the scope of this invention. One type of covalent modification included within the scope of this invention is a target polypeptide fragment. Target polypeptide fragments having up to about 40 amino acid residues may be conveniently prepared by chemical synthesis, or by enzymatic or chemical cleavage of the full-length target polypeptide or variant target polypeptide. Other types of covalent modifications of the target polypeptide or fragments thereof are introduced into the molecule by reacting specific amino acid residues of the target polypeptide or fragments thereof with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues.

Cysteinyl residues most commonly are reacted with N-haloacetates (and corresponding amines), such as chloroacetic acid or chloroacetamide, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteinyl residues also are derivatized by reaction with bromotrifluoroacetone, N-bromo-p- (5-imidazole)propionic acid, chloroacetyliophosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide, Other modifications include hydroxylation of proline and 2-chloromercuri-4-nitrophenoalkanes or chloro-7-nitrobenzo-2-oxa-1,3-diazole. Histidyl residues are derivatized by reaction with diethyl pyrocarbonate at pH 5.5-7.0 because this agent is relatively specific for the histidyl side chain. Parabromophenacyl bromide also is useful; the reaction is preferably performed in 0.1 M sodium cacodylate at pH 6.0.

Lysyl and amino terminal residues are reacted with succinic or other carboxylic acid anhydrides. Derivatization with these agents has the effect of reversing the charge of the lysyl residues. Other suitable reagents for derivatizing α-amino-containing residues include imidoesters such as methyl picolinimidate; pyridoxal phosphate; pyridoxal; arginine epsilon-amino group. The tri-peptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except glycine, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine residue. The tri-peptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine residue.

Glycosylation of polypeptides is typically either N-linked or O-linked. N-linked refers to the attachment of the carbohydrate moiety to the asparagine residue. The tri-peptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine residue. Thus, the presence of either of these tri-peptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of the carbohydrate moiety to the serine or threonine. Thus, the presence of either of these tri-peptide sequences in a polypeptide creates a potential glycosylation site.

A deglycosylation site to the target polypeptide is conveniently accomplished by altering the amino acid sequence such that it contains one or more of the above-described tri-peptide sequences (or N-linked glycosylation sites). The alteration may also be made by the addition of, or substitution by, one or more serine or threonine residues to the native target polypeptide sequence (for O-linked glycosylation sites). For ease, the target polypeptide amino...
acid sequence is preferably altered through changes at the DNA level, particularly by mutating the DNA encoding the target polypeptide at selected bases such that codons are generated that will translate into the desired amino acids. The DNA mutation(s) may be made using methods described above under the heading of "Amino Acid Sequence Variants of Target Polypeptide".

Another means of increasing the number of carbohydrate moieties on the target polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. These procedures are advantageous in that they do not require production of the polypeptide in a host cell that has glycosylation capabilities for N- and O-linked glycosylation. Depending on the coupling mode used, the engager may be attached to (a) arginine and histidine, (b) free carboxyl groups, (c) free sulfhydryl groups such as those of cysteine, (d) free hydroxyl groups such as those of serine, threonine, or hydroxyproline, (e) aromatic residues such as those of phenylalanine, tyrosine, or tryptophan, or (f) the amine group of glutamine. These methods are described in WO 87/05330 published Sept. 11, 1987, and in Apinl and Wiston (CRC Crit. Rev. Biochem., pp. 259-306 [1981]).

Removal of carbohydrate moieties present on the native target polypeptide may be accomplished chemically or enzymatically. Chemical deglycosylation requires exposure of the polypeptide to the appropriate trichloroacetic acid, or an equivalent compound. This treatment results in the cleavage of most or all sugars except the linking sugar (N-acetylglucosamine or N-acetylgalactosamine), while leaving the polypeptide intact. Chemical deglycosylation is described by Hikimuddin et al. (Arch. Biochem. Biophys., 259:52 [1987]) and by Edge et al. (Anal. Biochem., 118:131 [1981]). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al. (Methods Enzymol. 138:350 [1987]).

Glycosylation at potential glycosylation sites may be prevented by the use of the compound tunicamycin as described by Duskin et al. (J. Biol. Chem., 257:3105 [1982]). Tunicamycin blocks the formation of protein-N-glycoside linkages.

Another type of covalent modification of the target polypeptide comprises linking the target polypeptide to various nonproteinaceous polymers, e.g. polyethylene glycol, polypropylene glycol or polyoxyalkylene, in the manner set forth in U.S. Pat. Nos. 4,640,835; 4,496,689; 4,301,144; 4,607,417; 4,791,192 or 4,719,337.

The target polypeptide also may be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization (for example, hydroxyethylcellulose or gelatin-microcapsules and poly[methylmethacrylate]microcapsules, respectively), in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules), or in macromolecules. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th Edition, Osol, A., Ed., (1980).

Target polypeptide preparations are also useful in generating antibodies, for screening or binding partners, as standards in assays for the target polypeptide (e.g. by labeling the target polypeptide for use as a standard in a radioluminescence assay, enzyme-linked immunosorbent assay, or radioimmunoassay), in affinity purification techniques, and in competitive-type receptor binding assays when labeled with radioiodine, enzymes, fluorophores, spin labels, and the like.

Since it is often difficult to predict in advance the characteristics of a variant target polypeptide, it will be appreciated that some screening of the recovered variant will be needed to select the optimal variant. For example, a change in this immunological character of the target polypeptide molecule, such as affinity for a given antigen or antibody, is measured by a competitive-type immunoassay. The variant is assayed for changes in the suppression or enhancement of its activity bad comparison to the activity observed for the target polypeptide in the same assay. Other potential modifications of protein or polypeptide properties such as redox or thermal stability, hydrophobicity, proteolytic degradation, stability in recombinant cell culture, in plasma, or the tendency to aggregate with carriers or into multimers are assayed by methods well known in the art. Diagnostic and Related Uses of the Antibodies

The antibodies of this invention are useful in diagnostic assays for antigen expression in specific cells or tissues. The antibodies are detectably labeled and/or are immobilized on an insoluble matrix.

The antibodies of this invention find further use for the affinity purification of the antigen from recombinant cell culture or natural sources. Suitable diagnostic assays for the antigen and its antibodies depend on the particular antigen or antibody. Generally, such assays include competitive and sandwich assays, and steric inhibition assays. Competitive and sandwich methods involve a phase-separation step as an integral part of the method while steric inhibition assays are conducted in a single reaction mixture. Fundamentally, the same procedures are used for the assay of the antigen and for substances that bind the antigen, although certain methods will be favored depending upon the molecular weight of the substance being assayed. Therefore, the substance to be tested is referred to herein as an analyte, irrespective of its substance being assayed. Therefore, the substance to be tested is referred to herein as an analyte, irrespective of its

Analytical methods for the antigen or its antibodies all use one or more of the following reagents: labeled analyte analogue, immobilized analyte analogue, labeled binding partner, immobilized binding partner, and coupling agents. The labeled reagents also are known as "tracers."

The labeled used (and this is also useful to label antigen nucleic acid for use as a probe) is any detectable functionality that does not interfere with the binding of analyte and its binding partner. Numerous labels are known for use in immunoassay, examples including moieties that may be detected directly, such as fluorochrome, chemiluminescent, and radioactive labels, as well as moieties, such as enzymes, that must be reacted or derivatized to be detected. Examples of such labels include the radioisotopes 32P, 14C, 125I, 3H, and 131I, fluorophores such as rare earth chelates or fluorescent and its derivatives, rhodamine and its derivatives, dansyl, umbellifere, isocerberase, e.g., dicyflucres and bacterial luciferase (U.S. Pat. No. 4,737,450), luciferase, 2,3-dihydrothiaphtalazinediones, horseradish peroxidase (HRP), alkaline phosphatase, β-galactosidase, glucoamylase, lysozyme, saccharide oxidases, e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase, heterocyclic oxidases such as uricase and xanthine oxidase, coupled with an enzyme that employs hydroxylamine to oxidize a dye precursor such as HRP, lactoperoxidase, or microperoxidase, biotin/avidin, spin labels, bacteriophage labels, stable free radicals, and the like.

Conventional methods are available to bind these labels covalently to proteins or polypeptides. For instance, coupling agents such as dialdehydes, carbodiimides,
35

40

45

50

55

60
dimaleimides, bis-vidimidates, bis-diazotized benzidine, and
the like may be used to tag the antibodies with the above-des
scribed fluorescent, chemiluminescent, and enzyme labels. See, for example, U.S. Pat. No. 3,940,475
such as horseradish peroxidase and alkaline phosphatase.

The conjugation of such label, including the enzymes, to the
antibody is a standard manipulative procedure for one of
ordinary skill in immunoassay techniques. See, for example,
O’Sullivan et al., “Methods for the Preparation of Enzyme-
enzyme Antibody Conjugates for Use in Enzyme Immunoassay,” in
pp.147–166. Such binding methods are suitable for use with
the antibodies and polypeptides of this invention.

Immobilization of reagents is required for certain assay
methods. Immobilization entails separating the binding par-
ter from any analyte that remains free in solution in this
conventionally is accomplished by either insolubilizing the
binding partner or analyte analogue before the assay
proceeds, as by adsorption to a water-insoluble matrix or
surface (Bemich et al., U.S. Pat. No. 3,720,760), by cova-

coupling (for example, using glutaraldehyde cross-

This separation is accomplished by decanting (where the
binding partner was preinsolubilized) or by centrifuging
(when the binding partner was precipitated after the com-

petitive reaction). The amount of test sample analyte is
inversely proportional to the amount of bound tracer as
measured by the amount of marker substance. Dose-

response curves with known amounts of analyte are pre-
pared and compared with the test results to quantitatively
determine the amount of analyte present in the test sample.
These assays are called ELISA systems when enzymes are
used as the detectable markers.

Another species of competitive assay, called a “homoge-

eous” assay, does not require a phase separation. Here, a
conjugate of an enzyme with the analyte is prepared and
used such that when anti-analyte binds to the analyte the
presence of the anti-analyte modifies the enzyme activity. In
this case, the antigen or its immunologically active frag-
ments are conjugated with a bifunctional organic bridge to
an enzyme such as peroxidase. Conjugates are selected for
use with antibody so that binding of the antibody inhibits or
potentiates the enzyme activity of the label. This method per-
se is widely practiced under the name of EMIT.

Steric conjugates are used in steric hindrance methods for
homogeneous assay. These conjugates are synthesized by
covalently linking a low-molecular-weight hapten to a small
analyte so that antibody to hapten substantially is unable to
bind the conjugate at the same time as anti-analyte. Under
this assay procedure the analyte present in the test sample
will bind anti-analyte, thereby allowing anti-hapten to bind
the conjugate, resulting in a change in the character of the
conjugate hapten, e.g., a change in fluorescence when the
hapten is a fluorophore.

Sandwich assays particularly are useful for the determi-
nation of antigen or antibodies. In sequential sandwich
assays an immobilized binding partner is used to adsorb test
sample sample analyte, the test sample is removed as by washing,
the bound analyte is used to adsorb labeled binding partner,
and bound material is then separated from residual tracer.
The amount of bound tracer is directly proportional to test
sample analyte. In “simultaneous” sandwich assays the test
sample is not separated before adding the labeled binding
partner. A sequential sandwich assay using an anti-antigen
monoclonal antibody as one antibody and a polyclonal
anti-antigen antibody as the other is useful in testing samples
for particular antigen activity.

The foregoing are merely exemplary diagnostic assays for
the import and humanized antibodies of this invention.

Other assay methods, known as competitive or sandwich
assays, are well established and widely used in the com-
mercial diagnostics industry.

Competitive assays rely on the ability of a tracer analogue
to compete with the test sample analyte for a limited number
of binding sites on a common binding partner. The binding
partner generally is immobilized before or after the com-
petition and then the tracer and analyte bound to the binding
partner are separated from the unbound tracer and analyte.

Competitive or sandwich assays are also used when the
immunologically active fragments of the analyte are not
available intact. Conjugates of such analyte fragments with
antibody are used to assay the analyte. This method is
known as the “sandwich immunologically active fragment
assay” and is described in detail in the Examples.

Cytotoxins

This invention is also directed to immunochemical deriva-
tives of the antibodies of this invention such as immuno-
conjugates of a monoclonal antibody and a cytotoxic moiety.
Antibodies which carry the appropriate effector functions,
such as with their constant domains, are also used to induce
lysis through the natural complement process, and to interact
with antibody dependent cytotoxic cells normally present.

For example, purified, sterile filtered antibodies are
optionally conjugated to a cytotoxic such as ricin for use in
AIDS therapy. U.S. patent application Ser. No. 07/350,895
illustrates methods for making and using immunotoxins for
the treatment of HIV infection. The methods of this
invention, for example, are suitable for obtaining humanized
antibodies for use as immunotoxins in use for AIDS therapy.

The cytotoxic moiety of the immunotoxin may be a
cytotoxic drug or an enzymatically active toxin of bacterial,
fungal, plant or animal origin, or an enzymatically active
drug fragment of such a toxin. Enzymatically active toxins and
fragments thereof used are diphtheria A chain, nonbinding
active fragments of diphtheria toxin, exotoxin A chain (from
Pseudomonas aeruginosa), ricin A chain, shiga A chain,
modocin A chain, alpha-sarcin, Auleria aurantiifolia proteins,
diamin proteins, Phyotoallergens proteins (PAPI, PAPI, and
PAP-S), momordica charantia inhibitor, curcin, curcin,
saponaria officinalis inhibitor, gelonin, mitogellin,
restristin, phenomycin, enzymcin and the trichothecenes.
In another embodiment, the conjugates are conjugated to
small molecule anticancer drugs such as cis-platin or 5FU.

Conjugates of the monoclonal antibody and such cytotoxic
moieties are made using a variety of bifunctional protein
coupling agents. Examples of such reagents are SPDP, IT,
bifunctional derivatives of imidoesters such as dimethyl

diamide HCl, active esters such as diacetimimidyl

suberate, aldehydes such as glutaraldehyde, bis-azido com-

pounds such as bisp(azidobenzoyl)hexanediane, bis-
diazonium derivatives such as bis-(p-diazoniumbenzoyl)-

carbodimide, dilucosanates such as tolylene 2,6-

diisocyanate and bis-active fluoride compounds such as

1,5-difluoro-2,4-dinitrobenzene. The lysing portion of a
toxin may be joined to the Fab fragment of the antibodies.

Immunotoxins can be made in a variety of ways, as
discussed herein. Commonly known crosslinking reagents
can be used to yield stable conjugates.

Advantageously, monoclonal antibodies specifically bind-

ning the domain of the antigen which is exposed on the
infected cell surface, are conjugated to ricin A chain. Most advantageously the ricin A chain is deglycosylated and produced through recombinant means. An advantageous method of making the ricin immunotoxin is described in Vitetta et al., Science 238:1098 (1987).

When used to kill infected human cells in vitro for diagnostic purposes, the conjugates will typically be added to the cell culture medium at a concentration of at least about 10 nM. The formulation and mode of administration for in vitro use are not critical. Aqueous formulations that are compatible with the culture or perfusion medium will normally be used. Cytotoxicity may be read by conventional techniques.

Cytotoxic radiopharmaceuticals for treating infected cells may be made by conjugating radioactive isotopes (e.g., I, Y, Pr) to the antibodies. Advantageously alpha particle-emitting isotopes are used. The term "cytotoxic moiety" as used herein is intended to include such isotopes.

In a preferred embodiment, ricin A chain is deglycosylated and produced without attached oligosaccharides, to decrease its clearance by irrelevant clearance mechanisms (e.g., the liver). In another embodiment, whole ricin (A chain plus B chain) is conjugated to antibody if the galactose binding property of B-chain can be blocked ("blocked ricin").

In a further embodiment toxin-conjugates are made with Fab or (Fab') fragments. Because of their relatively small size these fragments can better penetrate tissue to reach infected cells.

In another embodiment, fusogenic liposomes are filled with a cytotoxic drug and the liposomes are coated with antibodies specifically binding the particular antigen.

Antibody Dependent Cellular Cytotoxicity

Certain aspects of this invention involve antibodies which are (a) directed against a particular antigen and (b) belong to a subclass or isotype that is capable of mediating the lysis of cells in which the antibody molecule binds. More specifically, these antibodies should belong to a subclass or isotype that, upon complexing with cell surface proteins, activates complement and/or mediates antibody dependent cellular cytotoxicity (ADCC) by activating effector cells such as natural killer cells or macrophages.

Biological activity of antibodies is known to be determined, to a large extent, by the constant domains or Fc region of the antibody molecule (Uman and Benacerraf, *Textbook of Immunology*, 2nd Edition, Williams & Wilkins, p. 218 (1984)). This includes their ability to activate complement and to mediate antibody-dependent cellular cytotoxicity (ADCC) as effected by leukocytes. Antibodies of different classes and subclasses differ in this respect, as do antibodies from the same subclass but different species; according to the present invention, antibodies of those classes having the desired biological activity are prepared.

Preparation of these antibodies involves the selection of antibody constant domains or their incorporation in the humanized antibody by known technique. For example, mouse immunoglobulins of the IgG3 and IgG2a class are capable of activating complement upon binding to the target cells which express the cognate antigen, and therefore humanized antibodies which incorporate IgG3 and IgG2a effector functions are desirable for certain therapeutic applications.

In general, mouse antibodies of the IgG2a and IgG3 subclasses and occasionally IgG1 can mediate ADCC, and antibodies of the IgG3, IgG2a, and IgM subclasses bind and activate complement. Complement activation generally requires the binding of at least two IgG molecules in close proximity on the target cell. However, the binding of only one IgM molecule activates complement.

The ability of any particular antibody to mediate lysis of the target cell by complement activation and/or ADCC can be assayed. The cells of interest are grown and labeled in vitro; the antibody is added to the cell culture in combination with either serum complement or immune cells which may be activated by the antigen antibody complexes. Cytolysis of the target cells is detected by the release of label from the lysed cells. In fact, antibodies can be screened using the patient's own serum as a source of complement and/or immune cells. The antibody that is capable of activating complement or mediating ADCC in the in vitro test can then be used therapeutically in that particular patient.

This invention specifically encompasses consensus Fc antibody domains prepared and used according to the teachings of this invention.

Therapeutic and Other Uses of the Antibodies

When used in vivo for therapy, the antibodies of the subject invention are administered to the patient in therapeutically effective amounts (i.e. amounts that have desired therapeutic effect). They will normally be administered parenterally. The dose and dosage regimen will depend upon the degree of the infection, the characteristics of the particular antibody or immunotoxin used, e.g., its therapeutic index, the patient, and the patient's history. Advantageously the antibody or immunotoxin is administered continuously over a period of 1-2 weeks, intravenously in order to treat cells in the vasculature and subcutaneously and intraperitoneally to treat regional lymph nodes. Optionally, the administration is made during the course of adjunct therapy such as combined cycles of radiation, chemotherapeutic treatment, or administration of tumor necrosis factor, interferon or other cytotoxic or immunomodulatory agent.

For parenteral administration the antibodies will be formulated in a unit dosage injectable form (solution, suspension, emulsion) in association with a pharmaceutically acceptable parenteral vehicle. Such vehicles are inherently nontoxic, and non-therapeutic. Examples of such vehicles are water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Nonaqueous vehicles such as fixed oils and ethyl oleate can also be used. Liposomes may be used as carriers. The vehicle may contain minor amounts of additives such as substances that enhance isotonicity and chemical stability, e.g., buffers and preservatives. The antibodies will typically be formulated in such vehicles at concentrations of about 1 mg/ml to 10 mg/ml.

Use of IgM antibodies may be preferred for certain applications, however IgG molecules by being smaller may be more able than IgM molecules to localize to certain types of infected cells.

There is evidence that complement activation in vivo leads to a variety of biological effects, including the induction of an inflammatory response and the activation of macrophages (Uman and Benacerraf, *Textbook of Immunology*, 2nd Edition, Williams & Wilkins, p. 218 (1984)). The increased vasoconstriction accompanying inflammation may increase the ability of various agents to localize in infected cells. Therefore, antigen-antibody combinations of the type specified by this invention can be used therapeutically in many ways. Additionally, purified antigens (Hakomori, *Ann Rev. Immunol* 2:103 (1984)) or antibody-idiotype antibodies (Nepom et al., *Proc. Natl Acad. Sci. USA* 81:268 (1985); Koprowski et al., *Proc. Natl Acad. Sci. USA* 81:216 (1984)) relating to such antigens could be used to induce an active immune response in human patients. Such a response includes the formation of antibodies capable of activating human complement and mediating ADCC and by such mechanisms cause infected cell destruction.
Optionally, the antibodies of this invention are useful in passively immunizing patients, as exemplified by the administration of humanized anti-HIV antibodies.

The antibody compositions used in therapy are formulated and dosages established in a fashion consistent with good medical practice taking into account the disorder to be treated, the condition of the individual patient, the site of delivery of the composition, the method of administration and other factors known to practitioners. The antibody compositions are prepared for administration according to the description of preparation of polypeptides for administration, infra.

Deposit of Materials

As described above, cultures of the muMAb4D5 have been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, Va., U.S.A. (ATCC).

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of viable cultures for 30 years from the date of the deposit. The organisms will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the cultures to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner’s rules pursuant thereto (including 37 CFR $1.12 with particular reference to 886 OG 638).

In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 25(4) EPC)

The assignee of the present application has agreed that if the cultures on deposit should die or be lost or destroyed when cultivated under suitable conditions, they will be promptly replaced on notification with a viable specimen of the same culture. Availability of the deposited strain is not to be construed as a license to practice the invention in countries for which the deposit was not made, even if the deposit was received by an expert nominated by the inventor of the deposit.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the constructs deposited, since the deposited embodiments are intended to illustrate only certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that they represent. Further, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

It is understood that the application of the teachings of the present invention to a specific problem or situation will be within the capabilities of one having ordinary skill in the art in light of the teachings contained herein. Examples of the products of the present invention and representative processes for their isolation, use, and manufacture are shown below, but should not be construed to limit the invention.

EXAMPLES

Example 1

Humanization of muMAb4D5

Here we report the chimerization of muMAb4D5 (chMAb4D5) and the rapid and simultaneous humanization of heavy (VH) and light (VL) chain variable region genes using a novel “gene conversion mutagenesis” strategy. Eight humanized variants (humMAb4D5) were constructed to probe the importance of several FR residues identified by our molecular modeling or previously proposed to be critical to the conformation of particular CDRs (see Chothia, C. & Lesk, A. M., J. Mol. Biol. 196:901–917 (1987); Chothia, C. et al., Nature 342:877–883 (1989); Tramontano, A. et al., J. Mol. Biol. 215:175–182 (1990)). Efficient transient expression of humanized variants in non-myeloma cells allowed us to rapidly investigate the relationship between binding affinity for p185HER2 ECD and anti-proliferative activity against p185HER2 overexpressing carcinoma cells.

Materials and Methods

Cloning of Variable Region Genes. The muMAb4D5 VH and VL genes were isolated by polymerase chain reaction (PCR) amplification of mRNA from the corresponding hybridoma (Fendly, B. M. et al., Cancer Res. 50:1550–1558 (1990)) as described by Orlando et al. (Orlando, R. et al., Proc. Natl. Acad. Sci. USA 86:3833–3837 (1990)). Amino terminal sequencing of muMAb4D5 VH and VL was used to design the sense strand PCR primers, whereas the anti-sense PCR primers were based upon consensus sequences of murine framework residues (Orlando, R. et al., Proc. Natl. Acad. Sci. USA 80:3833–3837 (1989); Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) incorporating restriction sites for directional cloning shown by underlining and listed after the sequences: VH sense, 5’-TCCGAGATCCAGCTGACCAGTCTCCA-3’ (SEQ. ID NO. 7), EcoRI; VH anti-sense, 5’-GGTTCAGCTCAAGCTTTCC-3’ (SEQ. ID NO. 8), Asp718; VL sense, 5’-AAGTSMARCTGCAGSAGTCWGG-3’ (SEQ. ID NO. 9), PstI; and VL anti-sense, 5’-TOGGAGGACGGTACCHSCDCCGAA-3’ (SEQ. ID NO. 10), BstEII; where H=A or C, S=C or G, D=A or G or T, M=A or C or T, R=A or G, and W=A or T. The PCR products were cloned into pUC19 (Vieira, J. & Messing, J., Methods Enzymol. 153:3-11 (1987)) and five clones for each variable domain sequenced by the dideoxy method (Sanger, F. et al., Proc. Natl. Acad. Sci. USA 74:5463–5467 (1977)).

Molecular Modelling. Models for muMAb4D5 VH and VL domains were constructed separately from consensus coordinates based upon seven Fab structures from the Brookhaven protein data bank (entries 1FB4, 2RHE, 2MCP, 3PAB, 1FB1, 2HL1 and 1REI). The Fab fragment KI1 (Marquart, M. et al., J. Mol. Biol. 141:369–391 (1980)) was first chosen as a template for VH and VL domains and additional structures were then superimposed upon this structure using their main chain atom coordinates (INSIGHT program, Biosym Technologies). The distances from the template Ca to the analogous Ca in each of the superimposed structures was calculated for each residue position. If all (or nearly all) Cα-Cα distances for a given residue were


In humanizing muMAb4DS, consensus human sequences were first derived from the most abundant subclasses in the sequence compilation of Kabat et al. (Kabat, E. A. et al., *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Bethesda, Md., 1987)), namely V<sub>k</sub> subgroup I and V<sub>h</sub> group III, and a molecular model generated for these sequences using the methods described above. A structure for huMAb4DS was created by transferring the CDRs from the muMAb4DS model into the consensus human structure. All huMAb4DS variants contain human replacements of muMAb4DS residues at three positions within CDRs as defined by sequence variability (Kabat, E. A. et al., *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Bethesda, Md., 1987)) but not as defined by structural variability (Chothia, C. & Lesk, A. M., *J. Mol. Biol. 196:901-917* (1987)) and designated CDR1 K24R, CDR2 P54I, and CDR3 T56N. Differences between muMAb4DS and the human consensus framework residues (Fig. 1) were individually modeled to investigate their possible influence on CDR conformation and/or binding to the p185<sup>HER2</sup> ECD.

Construction of Chimeric Genes Genes encoding chMAb4DS light and heavy chains were separately administered in previously described phagemid vectors containing the human cytoplasmic enhancer and promoter, a 5' intron, and SV40 polyadenylation signal (Gorman, C. M. et al., *DNA & Prot. Engin. Tech. 3:23-10* (1990)). Briefly, gene segments encoding muMAb4DS V<sub>k</sub> (Fig. 1A) and REI human V<sub>k</sub> light chain C<sub>Y</sub> (Palm, W. & Hilschmann, N. Z., *Physiol. Chem. 356:167-191* (1975)) were precisely joined as were genes for muMAb4DS V<sub>h</sub> (Fig. 1B) and human y1 constant region (Capon, D. J. et al., *Nature 337:525-531* (1989)) by simple subcloning (Boyle, A, in *Current Protocols in Molecular Biology*, Chapter 3 (F. A. Ausubel et al., eds., Greene Publishing & Wiley-Interscience, New York, 1990)) and site-directed mutagenesis (Carter, P., in *Molecular Biology, Chapter 3* (IRL Press, Oxford, UK 1991); and in Wells, J. A. et al., *Pharm. Trans. R. Soc. Lond. A 317:415-423* (1986)). Resultant clones containing both huV<sub>k</sub> and huV<sub>h</sub> genes were identified by nucleotide sequencing (Sanger, F. et al., *Proc. Natl. Acad. Sci. USA 74:5463-5467* (1977) and designated pAK2. Additional humanized variants were generated by site-directed mutagenesis (Carter, P., in *Molecular Biology, Chapter 3* (IRL Press, Oxford, UK 1991)). The muMAb4DS V<sub>k</sub> and V<sub>h</sub> gene segments in the transient expression vectors described above were then precisely replaced with their humanized versions. Expression and Purification of MAb4DS Variants. Appropriate MAb4DS light and heavy chain cDNA expression
vectors were co-transfected into an adenovirus transformed human embryonic kidney cell line, 293 (Graham, F. L. et al., J. Gen. Virol. 36:59–72 (1977)) using a high efficiency procedure (Gorman, C. M. et al., DNA & Prot. Engin. Tech. 2:3–10 (1990); Gorman, C., in DNA Cloning, vol II, pp 143–190 (D. M. Glover, ed., IRL Press, Oxford, UK 1985)). Media were harvested daily for up to 5 days and the cells re-fed with serum free media. Antibodies were recovered from the media and affinity purified on protein A sepharose CL-4B (Pharmacia) as described by the manufacturer. The eluted antibody was buffer-exchanged into phosphate-buffered saline by G25 gel filtration, concentrated by ultrafiltration (Centriprep-30 or Centricon-100, Amicon), sterile-filtered (Milliex-GV, Millipore) and stored at 4°C. The concentration of antibody was determined by using both total immunoglobulin and antigen binding ELISAs. The standard used was huMAb4D5-5, whose concentration had been determined by amino acid composition analysis.

Affinity Measurements. The antigen binding affinity of MAb4D5 variants was determined using a secreted form of the p185HER2 ECD prepared as described in Bendly, B. M. et al., Cancer Res. 50:1558–1561 (1990) using saturating MAb4DS concentrations.

Results

Humanization of muMAb4D5. The muMAb4DS Vg and Vh gene segments were first cloned by PCR and sequenced (Fig. 1). The variable genes were then simultaneously humanized by gene conversion mutagenesis using pre-synthesized oligonucleotides (Fig. 2). A 311-mer oligonucleotide containing 39 mismatches to the template directed 24 simultaneous amino acid changes required to humanize muMAb4DS Vg. Humanization of muMAb4DS Vh required 32 amino acid changes which were installed with a 361-mer containing 59 mismatches to the muMAb4DS template. Two out of the 8 clones sequenced precisely encode huMAb4DS-5, although one of these clones contained a single nucleotide imperfection. The 6 other clones were essentially humanized containing <3 nucleotide changes and <1 single nucleotide deletion per kilobase. Additional humanized variants (Table 3) were constructed by site-directed mutagenesis of huMAb4DS-5.

Expression levels of huMAb4DS variants were in the range of 75 µg/ml as judged by ELISA using immobilized p185HER2 ECD. Successive harvests of five 10 cm plates allowed 200 µg to 500 mg of each variant to be produced in a week. Antibodies affinity purified on protein A gave a single band on a Coomassie blue stained SDS polyacrylamide gel of mobility consistent with the expected Mr, of ~150 kDa. Electrophoresis under reducing conditions gave 2 bands consistent with the expected Mr, of free heavy (48 kDa) and light (23 kDa) chains (not shown). Amino terminal sequence analysis (16-cycles) gave the mixed sequence expected (Fig. 1) from an equimolar combination of light and heavy chains (not shown). huMAb4DS Variants. In general, the FR residues were chosen from consensus human sequences (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) but an arginine occupies this position in the muMAb4DS k light chain. The side chain of residue 66 is usually a glycine in human and murine k chain sequences (Kabat, E. A. et al., Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md., 1987)) but an arginine occupies this position in the muMAb4DS k light chain. The side chain of residue 66 is likely to affect the conformation of Vh-CDR1 and Vh-CDR2 and the hairpin turn at 68–69 (Fig. 4). Consistent with the importance of this residue, the mutation Vh-G66R (huMAb4D5-3–muMAb4DS-2) increases the affinity for the p185HER2 ECD by 4-fold with a concomitant increase in anti-proliferative activity.

From molecular modeling it appears that the tyrosyl side chain of muMAb4DS Vh residue 66 may either stabilize the conformation of Vh-CDR3 or provide an interaction at the
V_{i}-V_{j} interface. The latter function may be dependent upon the 
occurrence of V_{j} Y102. In the context of huMAb4D5-5 the 
mutations V_{2} E55Y (huMAb4D5-6) and V_{3} V102Y (huMAb4D5-7) 
individually increase the affinity for p185HER2 ECD by 5-fold and 2-fold respectively, whereas 
together (huMAb4D5-8) they increase the affinity by 
11-fold. This is consistent with either proposed role of V_{2} 
Y55 and V_{3} Y102.

Secondary Immune Function of huMAb4D5-8. 
MuMAb4D5 inhibits the growth of human breast tumor 
cells which overexpress p185HER2 (Hudziak, R. M. et al., 
Molec. Cell. Biol. 9:1165-1172 (1989)). The antibody, 
however, does not offer the possibility of direct tumor 
cytotoxic effects. This possibility does arise in 
huMAb4D5-8 as a result of its high affinity (K_{D}=0.1 \mu M) and 
its human IgG1 subtype. Table 4 compares the ADCC 
mediated by huMAb4D5-8 with muMAb4D5 on a normal 
human lung epithelial cell line, WI-38, which expresses a low level of 
p185HER2 and on SK-BR 3, which expresses a high level of 
p185HER2. The results demonstrate that: (1) huMAb4D5 
has a greatly enhanced ability to carry out ADCC as 
compared with its murine parent; and (2) that this activity may 
be selective for cell types which overexpress p185HER2.

Discussion 
MuMAb4D5 is potentially useful for human therapy since 
it is cytostatic towards human breast and ovarian tumor lines 
overexpressing the HER2-encoded p185HER2 receptor-like 
tyrosine kinase. Since both breast and ovarian carcinomas are chronic diseases it is anticipated that the optimal 
MAb4D5 variant molecule for therapy will have low immuno-
genicity and will be cytotoxic rather than solely cytostatic in effect. Humanization of muMAb4D5 should accomplish 
these goals. We have identified 5 different huMAb4D5 
variants which bind tightly to p185HER2 ECD (K_{D}=1 \mu M) and 
which have significant anti-proliferative activity (Table 3). 
Furthermore huMAb4D5-8 but not muMAb4D5 mediates 
ADCC against human tumor cell lines overexpressing 
p185HER2 in the presence of human effector cells (Table 4) 
as anticipated for a human IgG isotype (Brüggenmann, M. et al., 
Exp. Med. 166:1351-1361 (1987); Riechmann, L. et al., 
Nature 332:323-327 (1988)).

Rapid humanization of huMAb4D5 was facilitated by the 
gene conversion mutagenesis strategy developed here using 
long preassembled oligonucleotides. This method requires 
less than half the amount of synthetic DNA as does total 
gene synthesis and does not require convenient restriction 
sites in the target DNA. Our method appears to be simpler 
and more reliable than a variant protocol recently reported 
(Rostapshov, V. M. et al., FEBS Lett. 249:379-382 (1989)).

Table 3: The p185HER2 ECD binding affinity and anti-proliferative activities of MAb4D5 variants

<table>
<thead>
<tr>
<th>Variants</th>
<th>FR3</th>
<th>FR3</th>
<th>FR3</th>
<th>FR3</th>
<th>FR3</th>
<th>FR2</th>
<th>FR3</th>
<th>nM</th>
<th>Relative cell proliferation</th>
</tr>
</thead>
<tbody>
<tr>
<td>huMAb4D5-1</td>
<td>R</td>
<td>D</td>
<td>L</td>
<td>A</td>
<td>V</td>
<td>E</td>
<td>G</td>
<td>25</td>
<td>102</td>
</tr>
<tr>
<td>huMAb4D5-2</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>V</td>
<td>E</td>
<td>G</td>
<td>4.7</td>
<td>101</td>
</tr>
<tr>
<td>huMAb4D5-3</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>V</td>
<td>E</td>
<td>G</td>
<td>4.4</td>
<td>66</td>
</tr>
<tr>
<td>huMAb4D5-4</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>V</td>
<td>H</td>
<td>Arg</td>
<td>0.92</td>
<td>56</td>
</tr>
<tr>
<td>huMAb4D5-5</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>V</td>
<td>E</td>
<td>Arg</td>
<td>3.1</td>
<td>48</td>
</tr>
<tr>
<td>huMAb4D5-6</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>V</td>
<td>Tyr</td>
<td>Arg</td>
<td>0.22</td>
<td>51</td>
</tr>
<tr>
<td>huMAb4D5-7</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>Tyr</td>
<td>E</td>
<td>Arg</td>
<td>0.62</td>
<td>53</td>
</tr>
<tr>
<td>huMAb4D5-8</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>Tyr</td>
<td>E</td>
<td>Arg</td>
<td>0.19</td>
<td>54</td>
</tr>
<tr>
<td>muMAb4D5-1</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>Tyr</td>
<td>E</td>
<td>Arg</td>
<td>0.30</td>
<td>37</td>
</tr>
</tbody>
</table>

*Human and murine residues are shown in one letter and three letter amino acid code respectively.

K_{D} values for the p185HER2 ECD were determined using the method of Friguet et al. (43) and 
the standard error of each estimate is \pm 10%.
TABLE 3-continued

<table>
<thead>
<tr>
<th>Residue</th>
<th>Residue</th>
<th>Relative cell proliferation</th>
</tr>
</thead>
<tbody>
<tr>
<td>71</td>
<td>73</td>
<td>75</td>
</tr>
<tr>
<td>FR3</td>
<td>FR3</td>
<td>FR3</td>
</tr>
<tr>
<td>102</td>
<td>55</td>
<td>66</td>
</tr>
<tr>
<td>CDR3</td>
<td>CDR2</td>
<td>FR3</td>
</tr>
</tbody>
</table>

*Proliferation of SK-BR-3 cells incubated for 96 hr with MAAb4DS variants shown as a percentage of the untreated control as described (Hudilask, R. M., et al., Molecular Cell. Biol. 9:1165-1173 (1989)). Data represent the maximal anti-proliferative effect for each variant (see FR3, 3A) calculated as the mean of triplicate determinations at a MAAb4DS concentration of 0.1 µg/ml. Data are all taken from the same experiment with an estimated standard error of ±15%.

TABLE 4

<table>
<thead>
<tr>
<th>Effector/Target</th>
<th>WE-38</th>
<th>SK-BR-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25:1</td>
<td>&lt;1.0</td>
<td>9.3</td>
</tr>
<tr>
<td>12.5:1</td>
<td>&lt;1.0</td>
<td>11.3</td>
</tr>
<tr>
<td>6.25:1</td>
<td>&lt;1.0</td>
<td>8.9</td>
</tr>
<tr>
<td>3.13:1</td>
<td>&lt;1.0</td>
<td>8.5</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25:1</td>
<td>&lt;1.0</td>
<td>1.3</td>
</tr>
<tr>
<td>12.5:1</td>
<td>&lt;1.0</td>
<td>1.7</td>
</tr>
<tr>
<td>6.25:1</td>
<td>&lt;1.0</td>
<td>2.2</td>
</tr>
<tr>
<td>3.13:1</td>
<td>&lt;1.0</td>
<td>0.8</td>
</tr>
</tbody>
</table>

* Sensitivity to ADCC of two human cell lines (WE-38, normal lung epithelial; and SK-BR-3, human breast tumor cell line) are compared. WE-38 expresses a low level of p180HER2 (0.6 pg per µg cell protein) and SK-BR-3 expresses a high level of p180HER2 (64 pg p180HER2 per µg cell protein), as determined by ELISA (Feinberg, et al., J. Biol. Med. 135:1-351 (1990)).

Example 2

Schematic Method for Humanizing an Antibody Sequence

This example illustrates one stepwise elaboration of the methods for creating a humanized sequence described above. It will be understood that not all of these steps are essential to the claimed invention, and that steps may be taken in different order.

1. Ascertain a consensus human variable domain amino acid sequence and prepare from it a consensus structural model.

2. Prepare model of import (the non-human domain to be humanized) variable domain sequences and note structural differences with respect to consensus human model.

3. Identify CDR sequences in human and in import, both by using Kabat (supra, 1987) and crystal structure criteria. If there is any difference in CDR identity from the different criteria, use of crystal structure definition of the CDR, but retain the Kabat residues as important framework residues to import.

4. Substitute import CDR sequences for human CDR sequences to obtain initial “humanized” sequence.

5. Compare import non-CDR variable domain sequence to the humanized sequence and note divergences.

6. Proceed through the following analysis for each amino acid residue where the import diverges from the humanized.

a. If the humanized residue represents a residue which is generally highly conserved across all species, use the residue in the humanized sequence. If the residue is not conserved across all species, proceed with the analysis described in 6b.

b. If the residue is not generally conserved across all species, ask if the residue is generally conserved in humans.

i. If the residue is generally conserved in humans but the import residue differs, examine the structural models of the import and human sequences and determine if the import residue would be likely to affect the binding or biological activity of the CDRs by considering 1) could it bind antigen directly and 2) could it affect the conformation of the CDR. If the conclusion is that an affect on the CDRs is likely, substitute the import residue. If the conclusion is that a CDR affect is unlikely, leave the humanized residue unchanged.

ii. If the residue is also not generally conserved in humans, examine the structural models of the import and human sequences and determine if the import residue would be likely to affect the binding or biological activity of the CDRs by considering 1) could it bind antigen directly and 2) could it affect the conformation of the CDR. If the conclusion is that an affect on the CDRs is likely, substitute the import residue. If the conclusion is that a CDR affect is unlikely, leave the humanized residue unchanged.

Example 2

Schematic Method for Humanizing an Antibody Sequence

This example illustrates one stepwise elaboration of the methods for creating a humanized sequence described above. It will be understood that not all of these steps are essential to the claimed invention, and that steps may be taken in different order.

1. Ascertain a consensus human variable domain amino acid sequence and prepare from it a consensus structural model.

2. Prepare model of import (the non-human domain to be humanized) variable domain sequences and note structural differences with respect to consensus human model.

3. Identify CDR sequences in human and in import, both by using Kabat (supra, 1987) and crystal structure criteria. If there is any difference in CDR identity from the different criteria, use of crystal structure definition of the CDR, but retain the Kabat residues as important framework residues to import.

4. Substitute import CDR sequences for human CDR sequences to obtain initial “humanized” sequence.

5. Compare import non-CDR variable domain sequence to the humanized sequence and note divergences.

6. Proceed through the following analysis for each amino acid residue where the import diverges from the humanized.

a. If the humanized residue represents a residue which is generally highly conserved across all species, use the residue in the humanized sequence. If the residue is not conserved across all species, proceed with the analysis described in 6b.

b. If the residue is not generally conserved across all species, ask if the residue is generally conserved in humans.

i. If the residue is generally conserved in humans but the import residue differs, examine the structural models of the import and human sequences and determine if the import residue would be likely to affect the binding or biological activity of the CDRs by considering 1) could it bind antigen directly and 2) could it affect the conformation of the CDR. If the conclusion is that an affect on the CDRs is likely, substitute the import residue. If the conclusion is that a CDR affect is unlikely, leave the humanized residue unchanged.

ii. If the residue is also not generally conserved in humans, examine the structural models of the import and human sequences and determine if the import residue would be likely to affect the binding or biological activity of the CDRs by considering 1) could it bind antigen directly and 2) could it affect the conformation of the CDR. If the conclusion is that an affect on the CDRs is likely, substitute the import residue. If the conclusion is that a CDR affect is unlikely, leave the humanized residue unchanged.
a) Examine the structural models of the import and human sequences and determine if the residue is exposed on the surface of the domain or is buried within. If the residue is exposed, use the residue in the humanized sequence. If the residue is buried, proceed to the next step.

(i) Examine the structural models of the import and human sequences and determine if the residue is likely to affect the V$_\alpha$-V$_\beta$ interface. Residues involved with the interface include: 34L, 36L, 38L, 43L, 33L, 36L, 35L, 37L, 41L, 43L, 39L, 43H, 45H, 47H, 60H, 91H, 93H, 95H, 96H, and 103H. If no effect is likely, use the residue in the humanized sequence. If some effect is likely, substitute the import residue.

7. Search the import sequence, the consensus sequence and the humanized sequence for glycosylation sites outside the CDRs, and determine if this glycosylation site is likely to have any effect on antigen binding and/or biological activity. If no effect is likely, use the human sequence at that site; if some effect is likely, eliminate the glycosylation site or use the import sequence at that site.

8. After completing the above analysis, determine the planned humanized sequence and prepare and test a sample. If the sample does not bind well to the target antigen, examine the particular residues listed below, regardless of the question of residue identity between the import and humanized residues.

a. Examine particular peripheral (non-CDR) variable domain residues which might interact directly with a macromolecular antigen, examine the particular residues listed below, including the following residues (where the * indicates residues which have been found to interact with antigen based on crystal structures):

i. Variable light domain: 36, 46, 49*, 63-70
ii. Variable heavy domain: 2, 47*, 68, 70, 73-76

b. Examine particular variable domain residues which could interact with, or otherwise affect, the conformation of variable domain and CDRs, including the following (not including CDR residues themselves, since it is assumed that, because the CDRs interact with one another, any residue in one CDR could potentially affect the conformation of another CDR residue (L-LIGHT, II-HEAVY, residues appearing in bold are indicated to be structurally important according to the Chothia et al., Nature 342:877 (1989), and residues appearing in italic were altered during humanization by Queen et al. (PDL), Proc. Natl. Acad. Sci. U.S.A. 86:10079 (1989) and Proc. Natl. Acad. Sci. U.S.A. 88:2869 (1991));

i. Variable light domain:
   a) CDR-1 (residues 24L-34L): 2L, 4L, 66L-69L, 71H
   b) CDR-2 (residues 50L-56L): 35L, 46L, 47L, 48L, 49L, 58L, 62L, 64L-66L, 71L, 73L
   c) CDR-3 (residues 89L-97L): 2L, 4L, 36L, 98L, 37H, 45H, 47H, 58H, 60H
ii. Variable heavy domain:
   a) CDR-1 (residues 26H-33H): 2H, 4H, 24H, 36H, 71H, 73H, 75H, 81H, 94H
   c) CDR-3 (residues 95H-102H): examine all residues as possible interaction partners with this loop, because this loop varies in size and conformation much more than the other CDRs.

9. If after step 8 the humanized variable domain still is lacking in desired binding, repeat step 8. In addition, re-investigate any human residues which might affect the V$_\alpha$-V$_\beta$ interface (but which would not directly affect CDR conformation). Additionally, evaluate the accessibility of non-CDR residues to solvent.

Example 3

Engineering a Humanized Bispecific F(ab')$_2$ Fragment

This example demonstrates the construction of a humanized bispecific antibody (BsF(ab')$_2$ v9) by separate directed chemical coupling in vitro. BsF(ab')$_2$ v1 (anti-CD3/anti-p185HER2) was demonstrated to retarget the cytotoxic activity of human CD3' CTL in vitro against the human breast tumor cell line, SK-BR-3, which overexpresses the p185HER2 product of the protooncogene HER2. This example demonstrates the minimalist humanization strategy of installing as few murine residues as possible into a human antibody in order to recruit antigen-binding affinity and biological properties comparable to that of the murine parent antibody. This strategy proved very successful for the anti-p185HER2 arm of BsF(ab')$_2$ v1. In contrast BsF(ab')$_2$ v1 binds to T cells via its anti-CD3 arm much less efficiently than does the chimeric BsF(ab')$_2$ which contains the variable heavy and light chains of the murine parent anti-CD3 antibody. Here we have constructed additional BsF(ab')$_2$ fragments containing variant anti-CD3 arms with selected murine residues restored in an attempt to improve antibody binding to T cells. One such variant, BsF(ab')$_2$ v9, was created by replacing six residues in the second hypervariable loop of the anti-CD3 heavy chain variable domain of BsF(ab')$_2$ v1 with their counterparts from the murine parent anti-CD3 antibody. BsF(ab')$_2$ v9 binds to T cells (Jurkat) much more efficiently than does BsF(ab')$_2$ v1 and almost as efficiently as the chimeric BsF(ab')$_2$. This improvement in the efficiency of T cell binding of the humanized BsF(ab')$_2$ is an important step in its development as a potential therapeutic agent for the treatment of p185HER2-overexpressing cancers.

Bispecific antibodies (BsAbs) with specificities for tumor-associated antigens and CDRs have often been used in preference to intact BsAbs in retargeted cellular cytotoxicity to avoid the risk of killing innocent bystander cells binding to the Tc region of the antibody. An additional advantage of BsF(ab')$_2$ over intact BsAbs is that they are generally much simpler to prepare free of contaminating monospecific molecules (reviewed by Songsivilai, S. and Lachmann, P. J., Clin. Exp. Immunol. 79: 315-321 (1990) and Nolan, O. and O'Kennedy, R., Biochim. Biophys. Acta 1040: 1-11 (1990)). BsF(ab')$_2$ fragments are traditionally constructed by directed chemical coupling of Fab' fragments obtained by limited proteolysis plus mild reduction of the parent rodent monoclonal Ab (Brennan, M. et al., Science 229, 81-83 (1985) and Glennie, M. J. et al., J. Immunol. 139: 2367-2375 (1987)). One such BsF(ab')$_2$ fragment (anti-glioma associated antigen/anti-CD3) was found to have clinical efficacy in glioma patients (Nitta, T. et al., Lancet 335: 368-371 (1990) and anotherBsF(ab')$_2$ (anti-iodum chelate/anti-carciinoembryonic antigen) allowed clinical imaging of colorectal carcinoma (Stückeley, D. R. et al., Antibody, Immunomed. Radiopharm. 2: 1-13 (1989)). Future BsF(ab')$_2$ destined for clinical applications are likely
to be constructed from antibodies which are either human or at least "humanized" (Riechmann, L. et al., Nature 332: 323–327 (1988) to reduce their immunogenicity (Hale, O. et al., Lancet i: 1394–1399 (1988)). Recently a facile route to a fully humanized BsF(ab')2 fragment designed for tumor immunotherapy has been demonstrated (Shalaby, M. R. et al., J. Exp. Med. 175: 217–225 (1992)). This approach involves separate Fc / expression of coupling each Fab' arm followed by traditional directed chemical coupling in vitro to form the BsF(ab')2. One arm of the BsF(ab')2 was a humanized version (Carter, P., et al., Proc. Natl. Acad. Sci. USA (1992a) and Carter, P. et al., Biopol Technology 10: 163–167 (1992b)) of the murine monoclonal Ab 415 which is directed against the p185HER2 product of the protooncogene HER2 (C-cerbB-2) (Fendly, B. M. et al. Cancer Res. 50: 1550–1558 (1990)). The humanization of the anti-CD3 variant 415 is shown in Example 1 of this application. The second arm was a minimally humanized anti-CD3 antibody (Shalaby et al. supra) which was created by install- ed the CDR loops from the variable domains of the murine parent monoclonal Ab UCHT1 (Beverley, P. C. L. and Callard, R. E., Eur. J. Immunol. 11: 320–334 (1981)) into the humanized anti-p185HER2 antibody. The BsF(ab')2 fragments containing the most potent humanized anti-CD3 variant (v1) was demonstrated by flow cytometry to bind specifically to a tumor target overexpressing p185HER2 and to human peripheral blood mononuclear cells carrying CD3. In addition, BsF(ab')2 v1 enhanced the cytotoxic effects of activated human CTL 4-fold against SK-BB-3 tumor cells overexpressing p185HER2. The example describes efforts to improve the antigen binding affinity of the humanized anti-CD3 arm by the judicious recruitment of a small number of additional murine residues into the minimally humanized anti-CD3 variable domains.

**Materials and Methods**

**Construction of Mutations in the Anti-CD3 Variable Region Genes**

The construction of genes encoding humanized anti-CD3 variant 1 (v1) variable light (Vλ) and heavy (VH) chain domains in plasmid pUC119 has been described (Shalaby et al. supra). Additional anti-CD3 variants were generated using an efficient site-directed mutagenesis method (Carter, P., et al., Biopol. Technology: a practical approach, (M. J. McPherson, Ed.), Chapter 1, IRL Press, Oxford, UK (1991)) using mismatched oligonucleotides which either install or remove amino acid in one letter code followed by the residue corresponding coding changes are denoted by the starting number of additional murine residues into the minimalistic-ly humanized anti-CD3 variable domains.

**Oligonucleotides used are listed below using lowercase to indicate the targeted mutations.**

- **HX11**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX12**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX13**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX14**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX15**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX16**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX17**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX18**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**HX19**: 5' CTATGATAATACACCATATGCACTTCACAG GATG

**OLLIGONUCLEOTIDES**

- HX11 installs a site for EcoRV (bold). Anti-CD3 variant v10 was constructed from v9 by site-directed mutagenesis using oligonucleotide ID13. Mutants were verified by dideoxy-molecule sequencing (Sanger, F. et al., Proc. Natl. Acad. Sci. USA 74: 5463–5467 (1977)).

**E. coli Expression of Fab' Fragments**

The expression plasmid, pAK19, for the co-secretion of light chain and heavy chain Fab' fragment of the most preferred humanized anti-p18SHER2 variant, HuMAb4D5-8, is described in Carter et al., 1992b, supra. Briefly, the Fab' expression unit is bacitracin with both chains under the transcriptional control of the phoA promoter. Genes encoding humanized Vλ and VH domains are precisely based on their 5' side to a gene segment encoding the heat-stable enterotoxin II signal sequence and on their 3' side to human k, Cb, and IgG1 Cγ1 constant domain genes, respectively. The Cα4 gene is immediately followed by a sequence encoding the hinge sequence CysAlaAla and followed by a bacteriophage λ ts, transcriptional terminator. Fab' expression plasmids for chimeric and humanized anti-CD3 variants (v1 to v4, Shalaby et al., supra; v6 to v12, this study) were created from pAK19 by precisely replacing anti-p18SHER2 Vλ and VH gene segments with those encoding murine and corresponding humanized variants of the anti-CD3 antibody, respectively, by sub-cloning and site-directed mutagenesis. The Fab' expression plasmid for the most potent humanized anti-CD3 variant identified in this study (v9) is designated pAK22. The anti-p18SHER2 Fab' fragment was secreted from E. coli K12 strain 218t2 containing plasmid pAK19 grown for 32 to 48 hr at 37°C in an aerated 10 liter fermentor. The final cell density was 120–150 OD600, and the titer of soluble and functional anti-p18SHER2 Fab' was 1–2 μg/ml as judged by antigen binding ELISA (Carter et al., 1992b, supra).

**Anti-CD3 Fab' variants were selected from E. coli containing corresponding expression plasmids using very similar fermentation protocols. The highest expression titers of chimeric and humanized anti-CD3 variants were 200 mg/liter and 700 mg/liter, respectively, as judged by total immunoglobulin ELISA.**

**Construction of BsF(ab')2 Fragments**

Fab' fragments were directly recovered from E. coli fermentation pastes in the free thiol form (Fab'-SH) by affinity purification on Streptococcal protein G at pH 5 in the presence of EDTA (Carter et al., 1992b, supra). The Fab' linked BsF(ab')2 fragments (anti-p18SHER2 Anti-CD3) were constructed by the procedure of Glennie et al. supra with the following modifications. Anti-p18SHER2 Fab'-SH in 100 mM Tris acetate, 5 mM EDTA (pH 5.0) was reacted with 0.1 vol. of 40 mM N,N',1,2-phenylenediamine (o-PDM) in dimethyl formamide for ~1.5 hr at 20°C. Excess o-PDM was removed by protein G purification of the Fab' maleimide derivative (Fab'-mal) followed by buffer exchange into 20 mM sodium acetate, 5 mM EDTA (pH 5.3) (coupling buffer) using centriprep-30 concentrators (Amicon). The total concentration of Fab' variants was estimated by reaction with 5,5'-dithiobis-(2-nitrobenzoic acid) as described by Creighton, T. E., Protein structure: a practical approach, (T. E. Creighton, Ed.), Chapter 7, IRL Press, Oxford, UK (1990). Equinomlar amounts of anti-p18SHER2 Fab'-mal (assuming quantitative reaction of Fab'-SH with o-PDM) and each anti-CD3 Fab'- SH variant were coupled together at a combined concentration of 1 to 2.5 mg/ml in the coupling buffer at 4°C. The coupling reaction was adjusted to 4 mM cysteine at pH 7.0 and incubated for 15 min at 20°C to
reduce any unwished disulfide-linked F(ab')2 formed. These reduction conditions are sufficient to reduce inter-heavy chain disulfide bonds with virtually no reduction of the disulfide link between light and heavy chains. Any free thiols generated were then blocked with 50 mM iodoacetamide. BsF(ab'), was isolated from the coupling reaction by Super-HR (Pharmacia) size exclusion chromatography (2.5 cm x 100 cm) in the presence of PBS. The BsF(ab'), samples were passed through a 0.2 mm filter flash frozen in liquid nitrogen and stored at -70° C.

Flow Cytometric Analysis of F(ab'), Binding to Jurkat Cells

The Jurkat human acute T cell leukemia cell line was purchased from the American Type Culture Collection (ATCC TIB 152) and grown as recommended by the ATCC. Aliquots of 10

 Jurkat cells were incubated with appropriate concentrations of BsF(ab'), (anti-p185ZR2 anti-CD3 variant) or control mono-specific anti-p185HER2 Fab', in PBS plus 0.1% (w/v) bovine serum albumin and 10 mM sodium azide for 45 min at 4° C. The cells were washed and then incubated with fluorescein-conjugated goat anti-human F(ab')2 (Organon Teknika, West Chester, Pa.) for 45 min at 4° C. Cells were washed and analyzed on a FACScan (Becton Dickinson and Co., Mountain View, Calif.). Cells (6 x 10^6) were acquired by list mode and gated by forward light scatter versus side light scatter excluding dead cells and debris.

Results

Design of Humanized Anti-CD3 Variants

The most potent humanized anti-CD3 variant previously identified, v1, differs from the murine parent antibody, UCHT1 at 19 out of 107 amino acid residues within V
c and at 37 out of 122 positions within V
d (Shalaby et al., supra 1992). Here we recruited back additional murine residues into anti-CD3 v1 in an attempt to improve the binding affinity for CD3. The strategy chosen was a compromise between minimizing both the number of additional murine residues recruited and the number of anti-CD3 variants to be analyzed. We focused our attention on a few CDR residues which were originally kept as human sequences in our minimalistic humanization regime. Thus human residues in V
c CDR2 of anti-CD3 v1 were replaced en bloc with their murine counterparts to generate anti-CD3 v9: TS75I/4SNpD155C3V447H (SEQ ID NO: 20) (FIG. 5). Similarly, the human residue E55 in V
d CDR2 of anti-CD3 v1 was replaced with histidine from the murine anti-CD3 antibody to generate anti-CD3 v11. In addition, V
c framework region (FR) residues 75 and 76 in anti-CD3 v1 were also replaced with their murine counterparts to create anti-CD3 v8: K75S:N76S. V
c residues 75 and 76 are located in a loop close to V
c CDR1 and CDR2 and therefore might influence antigen binding. Additional variants created by combining mutations at these three sites are described below.

Preparation of BsF(ab'), Fragments

Soluble and functional anti-p185HER2 and anti-CD3 Fab' fragments were recovered directly from corresponding E. coli fermentation broth with the single hinge with predominately in the free thiol form (75-100% Fab'-SH) by affinity purification on Streptococcal protein G at pH 5 in the presence of EDTA (Carter et al., 1992b, supra). Thioldirected BsF(ab'), fragments were then constructed by directed coupling using o-PDM as described by Glennie et al., supra. One arm was always the most potent humanized anti-p185HER2 variant, HuMAMDS-8 (Carter et al., 1992a, supra) and the other either a chimeric or humanized variant of the anti-CD3 antibody. Anti-p185HER2 Fab'-SH was reacted with o-PDM to form the maleimide derivative (Fab'-mal) and then coupled to the Fab'-SH for each anti-CD3 variant. F(ab')2 was then purified away from unreacted Fab' by size exclusion chromatography as shown for a representative preparation (BsF(ab'), v8) in data not shown. The F(ab')2 fragment represents ~5% of the total amount of antibody fragments (by mass) as judged by integration of the chromatograph peaks.

SDS-PAGE analysis of this BsF(ab'), v8 preparation under non-reducing conditions gave one major band with the expected mobility (Mr = 96 kD) as well as several very minor bands (data not shown). Amino-terminal sequence analysis of the major band after electrophoresing on to polyvinylidene difluoride membrane Matsudaira, P., J. Biol. Chem. 262:10035-10038 (1987) gave the expected mixed sequence from a stoichiometric 1:1 mixture of light and heavy chains (V
c/V
d: DE, V/E, Q/Q, M/L, TN, Q/E, S/S) expected for BsF(ab'),. The amino terminal region of both light chains are identical as are both heavy chains and correspond to consensus human FR sequences. We have previously demonstrated that F(ab'), constructed by directed chemical coupling carry both anti-p185HER2 and anti-CD3 antigen specificities (Shalaby et al., supra). The level of contamination of the BsF(ab'), with monospecific F(ab'), is likely to be very low since mock coupling reactions with both anti-p185HER2 Fab'-mal or anti-CD3 Fab'-SH alone did not yield detectable quantities of F(ab'),s. Furthermore the coupling reaction was subjected to a mild reduction step followed by alkylation to remove trace amounts of disulfide-linked Fab'-SH that might be present. SDS-PAGEd of the purified F(ab'), under reducing conditions gave two major bands with electrophoretic mobility and amino terminal sequence identical to free light chain and thioether-linked heavy chain dimers.

Scanning LASER densitometry of a o-PDM coupled F(ab'), preparation suggest that the minor species together represent ~10% of the protein. These minor contaminants were characterized by amino terminal sequence analysis and were tentatively identified on the basis of stoichiometry of light and heavy chain sequences and their electrophoretic mobility (data not shown). These data are consistent with the minor contaminating proteins in which the disulfide bond between light and heavy chains is missing in one or both arms, trace amounts of Fab' and heavy chain thioether-linked to light chain.

Binding of BsF(ab'), to Jurkat Cells

Binding of BsF(ab'), containing different anti-CD3 variants to Jurkat cells (human acute T cell leukemia) was investigated by flow cytometry (data not shown). BsF(ab')
v9 binds much more efficiently to Jurkat cells than does our starting molecule, BsF(ab'), v1, and almost as efficiently as the chimeric BsF(ab')
v2. Installation of additional murine residues into anti-CD3 v9 to create v10 (V
c K75S:N76S) and v12 (V
c K75S:N76S plus V
d E55H) did not further improve binding of corresponding BsF(ab'), to Jurkat cells. Nor did recruitment of these murine residues into anti-CD3 v1 improve Jurkat binding: V
c K75S:N76S (v6), V
c K75S:N76S (v7), V
d K75S:N76S (v8), V
d F55H (v11) (not shown) BsF(ab'), v9 was chosen for future study since it contains efficient variants in binding to Jurkat cells and contains fewest murine residues in the humanized anti-CD3 arm. A monospecific anti-p185HER2 Fab' (F(ab'),) did not show significant binding to Jurkat cells consistent with the interaction being mediated through the anti-CD3 arm.

Discussion

A minimalistic strategy was chosen to humanize the anti-p185HER2 (Carter et al., 1992a, supra) and anti-CD3
arms (Shalaby et al., supra) of the BsF(ab')

2 fragments in this study in an attempt to minimize the potential immunogenicity of the resulting humanized antibody in the clinic. Thus we tried to install the minimum number of murine CDR and FR residues into the context of consensus human variable domain sequences as required to recruit antigen-binding affinity and biological properties comparable to the murine parent antibody. Molecular modeling was used firstly to predict the murine FR residues which might be important to antigen binding and secondly to predict the murine CDR residues that might not be required. A small number of humanized variants were then constructed to test these predictions.

Our humanization strategy was very successful for the anti-p185

3 expressed antibody where one out of eight humanized variants (HuMAb4D5-8, IgG 1) was identified that bound the p185

4 isoform antigen ~3-fold more tightly than the parent murine antibody (Carter et al., 1992a, supra). HuMAb4D5-8 contains a total of five murine FR residues and nine murine CDR residues, including V

5 CDR2 residues 60-65, which were discarded in favor of human counterparts. In contrast, BsF (ab'), vl containing the most potent humanized anti-CD3 variant out of four originally constructed (Shalaby et al., supra) binds J6 cells with an affinity (K

6 ) of 140 nM which is ~70-fold weaker than that of the corresponding chimeric BsF(ab')

7 .

Here we have restored T cell binding of the humanized anti-CD3 close to that of the chimeric variant by replacing six murine residues in V

8 CDR2 with their murine counterparts: T58S:A60N:D61Q:S62K:V63I?G65D (anti-CD3 v9, FIG. 5). It appears more likely that these murine residues enhance antigen binding indirectly by influencing the conformation of residues in the N-terminal part of V

9 CDR2 rather than by directly contacting antigen. Finally, only N-terminal residues in V

10 CDR2 (50-58) have been found to contact antigen in one or more of eight crystallographic structures of antibody/antigen complexes (Kabat et al., supra; and Mian, I. S. et al., J. Mol. Biol. 217: 133-151 (1991), FIG. 5). Secondly, molecular modeling suggests that residues in the C-terminal part of V

11 CDR2 are at least partially buried (FIG. 5). BsF(ab')

12 binds to SK-RR-3 breast tumor cells with equal efficiency to BsF(ab'), vl and permits their direct recovery as Fab'-SH (Carter et al., 1992a, supra; and Glennie et al., supra). Traditional directed chemical coupling of thioether-linked BsF(ab'), vl has a 3-fold longer plasma residence time than BsF(ab'), vl linked by a single disulfide bond. Disulfide and thioether-linked chimeric BsF(ab'), were found to be indistinguishable in their efficiency of cell binding and in their retargeting of CTL cytotoxicity, which suggests that o-PDM directed coupling does not compromise binding of the BsF(ab')

13 to either antigen (not shown). Nevertheless the nature of the linkage appears not to be critical since a disulfide-linked BsF(ab')

14 murine anti-185

15/antibody was recently shown by others (Nishimura et al., Int. J. Cancer 50: 800-804 (1992) to have potent anti-tumor activity in nude mice. Our previous study (Shalaby et al., supra) together with this one and that of Nishimura, T. et al., supra improve the potential for using BsF(ab'), in targeted immunotherapy of p185

16 Overexpressing cancers in humans.

Example 4

Humanization of an Anti-CD18 Antibody

A murine antibody directed against the leukocyte adhesion receptor beta-chain (known as the H52 antibody) was humanized and several versions were created following the methods described above. FIGS 6A and 6B provide amino acid sequence comparisons for light chains and heavy chains of the murine and a preferred humanized antibody. FIG. 7A shows an amino acid sequence alignment of the sequences of thirteen humanized H52 variants (SEQ. ID. NOs 26-38); FIG. 7B shows an amino acid sequence alignment of two humanized H52 light chain variants (SEQ. ID. NOs 39-40). A specific light chain amino acid sequence may be paired with a selected heavy chain amino acid sequence to obtain antibody or antibody fragment molecules with desired binding characteristics.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(3) NUMBER OF SEQUENCES: 47

(1) INFORMATION FOR SEQ ID NO:1:

(4) SEQUENCE CHARACTERISTICS:

(a) STRAIN: 100 amino acids

(b) TYPE: Amino Acid
**(D) TOPOLOGY: Linear**

**[SEQUENCE DESCRIPTION: SEQ ID NO:1]**

Asp Ile Gin Met Thr Gln Ser Pro Ser Ser Ser Leu Ser Ala Ser Val 1 5 10 15
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gin Asp Val Asn 20 25 30
Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys 35 40 45
Leu Leu Ile Tyr Ser Ser Phe Leu Glu Ser Gly Val Pro Ser 50 55 60
Arg Phe Ser Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile 65 70 75
Ser Ser Leu Gln Pro Glu Phe Ala Thr Tyr Cys Gln Gln 80 85 90
His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Ala 95 100 105
Ile Lys Arg Thr 109

**[INFORMATION FOR SEQ ID NO:1]**

**(i) SEQUENCE CHARACTERISTICS:**
(A) LENGTH: 120 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

**[SEQUENCE DESCRIPTION: SEQ ID NO:2]**

Glu Val Gin Leu Val Gln Ser Gly Cys Gly Leu Val Gin Pro Gly 1 5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gin Asp Val Asn 20 25 30
Asp Thr Tyr Ile Ser Trp Val Arg Gln Ala Pro Gly Lys Gln Leu 35 40 45
Glu Trip Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr 50 55 60
Ala Asp Ser Val Lys Glu Arg Phe Thr Ile Ser Ala Asp Thr Ser 66 70 75
Lys Asp Thr Ala Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp 80 85 90
Thr Ala Val Tyr Cys Ser Arg Thr Gly Asp Gly Phe Tyr 95 100 105
Ala Met Asp Val Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser 110 115 120

**[INFORMATION FOR SEQ ID NO:2]**

**(i) SEQUENCE CHARACTERISTICS:**
(A) LENGTH: 120 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

**[SEQUENCE DESCRIPTION: SEQ ID NO:3]**

Asp Ile Gin Met Thr Gln Ser Pro Ser Ser Ser Leu Ser Ala Ser Val 1 5 10 15
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gin Asp Val Asn 20 25 30
Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys

**[INFORMATION FOR SEQ ID NO:3]**

**(i) SEQUENCE CHARACTERISTICS:**
(A) LENGTH: 109 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear
6,054,297

35 40 45

Leu Leu Ile Tyr Ala Ala Ser Ser Leu Glu Ser Gly Val Pro Ser

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile

Ser Ser Leu Gln Pro Gln Phe Ala Thr Tyr Cys Gln Gln

Tyr Asn Ser Leu Pro Tyr Thr Phe Gly Gly Thr Lys Val Glu

Ile Lys Arg Thr

---continued---

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 120 amino acids
   (B) TYPE: Amino Acid
   (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:I:

Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gin Pro Gly

Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser

Asp Tyr Ala Met Ser Trp Val Arg Gin Ala Pro Gly Lys Gly Leu

Glu Trp Val Ala Val Ile Ser Gin Ser Gin Thr Tyr Tyr Ala

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ser

Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp

Thr Ala Val Tyr Cys Ala Arg Asp Arg Gly Gly Ala Val Ser

Tyr Val Asp Val Trp Gln Gly Gly Thr Leu Val Thr Val Ser Ser

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 109 amino acids
   (B) TYPE: Amino Acid
   (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Ile Val Met Thr Ile Gln Ser His Lys Phe Met Ser Thr Ser Val

Gly Asp Arg Val Ser Ile Thr Cys Lys Ala Ser Gin Arg Val Asn

Thr Ala Val Ala Trp Tyr Gin Gin Lys Pro Gly His Ser Pro Lys

Leu Leu Ile Tyr Ser Ala Ser Phe Arg Tyr Thr Gly Val Pro Asp

Arg Phe Thr Gly Asn Arg Ser Gly Thr Asp Phe Thr Phe Thr Ile

Ser Ser Val Gin Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gin Gln

95 100 105

Ile Val Val Leu Val Ser Leu Thr Leu Ser Leu Met Asp Ser Val

Thr Arg Met Val Met Met Ser Phe Met Ser Thr Ser Val

Gly Thr Gly Gin Gin Ser Gin Arg Ala Leu Gln Pro Arg Val Asp

Tyr Asp Val Ile Thr Phe His Met Ser Thr Ser Val

Glu Thr Val Thr Val Ser Ser Ser Leu Leu Ile Met Thr Ser Met

Tyr Thr Ser Tyr Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Ser
(2) INFORMATION FOR SEQ ID NO:6:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 120 amino acids
       (B) TYPE: Amino Acid
       (D) TOPOLOGY: Linear

   (ii) SEQUENCE DESCRIPTION: SEQ ID NO:6:

       Glu Val Gin Leu Gin Gin Ser Gly Pro Glu Leu Val Lys Pro Gly
       1   5  10  15
       Ala Ser Leu Lys Leu Ser Cys Thr Ala Ser Gly Phe Asn Ile Lys
       20  25  30
       Asp Thr Tyr Ile His Thr Val Lys Gin Arg Pro Glu Gin Gly Leu
       35  40  45
       Glu Thr Ile Gly Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr
       50  55  60
       Asp Pro Lys Phe Gin Lys Ala Thr Ile Thr Ala Asp Thr Ser
       65  70  75
       Ser Asn Thr Ala Tyr Leu Val Ser Arg Leu Thr Ser Glu Asp
       80  85  90
       Thr Ala Val Tyr Tyr Cys Ser Arg Tyr Gly Arg Lys Phe Tyr
       95 100 105
       Ala Met Arg Tyr Ser Gly Gin Ala Ser Val Thr Val Ser
       110 115 120

(2) INFORMATION FOR SEQ ID NO:7:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 27 base pairs
       (B) TYPE: Nucleic Acid
       (C) STRANDEDNESS: Single
       (D) TOPOLOGY: Linear

   (ii) SEQUENCE DESCRIPTION: SEQ ID NO:7:

       TCCGATATCC
       AGCTGACCCA
       GTCTCCA

(2) INFORMATION FOR SEQ ID NO:8:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 31 base pairs
       (B) TYPE: Nucleic Acid
       (C) STRANDEDNESS: Single
       (D) TOPOLOGY: Linear

   (ii) SEQUENCE DESCRIPTION: SEQ ID NO:8:

       GTTTGATCTC
       CAGCTTGGTAC
       CCXSCDCCGA

(2) INFORMATION FOR SEQ ID NO:9:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 22 base pairs
       (B) TYPE: Nucleic Acid
       (C) STRANDEDNESS: Single
       (D) TOPOLOGY: Linear

   (ii) SEQUENCE DESCRIPTION: SEQ ID NO:9:

       AGGTSMARCT
       CGAGSAGTCU
       GG
(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
TGAGGAGACG GTGACCGTGG TCCCTTGGCC CCAG 34

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
CTGAGATAAAT CCTCTTACAG AGCCTATCTG CAAATG 36

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
CTGAGATAAAT CCTCTTACAG AGCCTATCTG CAAATG 36

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
CTGAGATAAAT CCTCTTACAG AGCCTATCTG CAAATG 36

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
CTTATAAAGG TGTTTCCACC TATAACCAGA AATTCAAGGA TCGTTACCG ATATCCGTAG ATAAA'KC 50

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
CTTACCTCC GCTTGAGCC TCTATACCTCC CGTCTGCATT CTGGAGTCCC 30
(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu
1  5  10  15
Gly Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Arg
20  25  30
Asn Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gln Thr Val Lys
35  40  45
Leu Leu Ile Tyr Thr Thr Ser Arg Leu His Ser Gln Val Pro Ser
50  55  60
Lys Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Leu Thr Ile
65  70  75
Ser Asn Leu Glu Gln Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln
80  85  90
Gly Asn Thr Leu Pro Trp Thr Phe Ala Gly Gly Thr Lys Leu Glu
95 100 105
Ile Lys
107

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Val
1  5  10  15
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Arg
20  25  30
Asn Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gln Ala Pro Lys
35  40  45
Leu Leu Ile Tyr Thr Ser Arg Leu Gln Ser Gln Val Ser Ser
50  55  60
Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile
65  70  75
Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Cys Gln Gln
80  85  90
Gly Asn Thr Leu Pro Trp Thr Phe Gly Gly Thr Lys Val Glu
95 100 105
Ile Lys
107

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
<table>
<thead>
<tr>
<th>6,054,297</th>
<th>75</th>
</tr>
</thead>
</table>

Asp Ile Gin Met Thr Glu Ser Pro Ser Ser Leu Ser Ala Ser Val
1 5 10 15

Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser
20 25 30

Asn Tyr Leu Ala Trp Tyr Gin Gin Lys Pro Gly Lys Ala Pro Lys
35 40 45

Leu Leu Ile Tyr Ala Ser Ser Leu Glu Ser Gly Val Pro Ser
50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile
65 70 75

Ser Ser Leu Gin Pro Glu Asp Phe Ala Thr Tyr Cys Gin Gin
80 85 90

Tyr Asn Ser Leu Pro Trp Thr Phe Gly Gin Thr Lys Val Glu
95 100 105

IIe Lys
107

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
Glu Val Gin Leu Gin Gin Ser Gly Pro Glu Leu Val Lys Pro Gly
1 5 10 15
Ala Ser Met Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser Thr Thr
20 25 30

Gly Tyr Thr Met Ser Thr Val Lys Gin Ser His Gly Lys Asn Leu
35 40 45

Glu Thr Met Gly Leu Ile Asn Pro Tyr Lys Gly Val Ser Thr Tyr
50 55 60

Asn Gin Lys Phe Lys Asp Lys Ala Thr Leu Thr Val Asp Lys Ser
65 70 75
Ser Ser Thr Ala Tyr Met Glu Leu Leu Ser Leu Thr Ser Glu Asp
80 85 90

Ser Ala Val Tyr Cys Ala Arg Ser Gly Tyr Gly Asp Ser
95 100 105

Asp Thr Tyr Phe Asp Val Trp Gly Ala Gly Thr Thr Val Thr Val
110 115 120

Ser Ser
122

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
Glu Val Gin Leu Val Glu Ser Gly Pro Gly Leu Val Lys Pro Gly
1 5 10 15

Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Sec Phe Thr
20 25 30

Gly Tyr Thr Met Asn Trp Val Arg Gin Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ala Leu Ile Asn Pro Tyr Lys Gly Val Ser Thr Tyr
59 60
Asn Glu Lys Phe Lys Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
65 78
Lys Asn Thr Ala Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80 85 90
Thr Ala Val Tyr Tyr Cys Ala Arg Ser Gly Tyr Tyr Asp Ser
95 100 105
Asp Trp Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val Thr Val
110 115 120
Ser Ser 122

(2) INFORMATION FOR SEQ ID NO:21:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
Glu Val Gin Leu Val Glu Ser Gly Gly Leu Val Glu Pro Gly
1 5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
20 25 30
Ser Tyr Ala Met Ser Trp Val Arg Glu Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ser Val Ile Ser Gly Asp Gly Ser Thr Tyr Tyr
50 55 60
Asp Ser Val Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser
65 70 75
Lys Asn Thr Leu Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80 85 90
Thr Ala Val Tyr Tyr Cys Ala Arg Gly Arg Val Gly Tyr Ser Leu
95 100 105
Ser Gly Leu Tyr Asp Tyr Trp Gly Gin Gly Thr Leu Val Thr Val
110 115 120
Ser Ser 122

(2) INFORMATION FOR SEQ ID NO:22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 454 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
Glu Val Gin Leu Val Glu Ser Gly Gly Leu Val Glu Pro Gly
1 5 10 15
Glu Tyr Thr Met Wle Thr Met Lys Gin Ser Cys Ser Lys Thr Ser
20 25 30
Glu Tyr Thr Met Wle Thr Met Lys Gin Ser Cys Ser Lys Ser
35 40 45
Glu Trp Ile Gly Phe Asn Pro Lys Asn Gly Gly Ser Ser
50 55 60
Asn Gin Arg Phe Met Asp Lys Ala Thr Leu Ala Val Asp Lys Ser
Thr Ser Thr Ala Tyr Met Glu Leu Arg Ser Leu Thr Ser Glu Asp
Ser Gly Ile Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
Thr Asp Val Arg Tyr Phe Asp Val Trp Gly Ala Gly Thr Thr Val
Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
Ala Pro Ser Ser Lys Ser Thr Ser Glu Gly Thr Ala Leu Gly
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Val Val Thr Val
Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn
His Lys Pro Ser Asn Thr Lys Val Asp Lys Val Glu Pro Lys
Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
Leu Leu Gln Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
Glu Gin Tyr Asn Ser Tyr Arg Val Val Ser Val Leu Thr Val
Leu His Glu Asp Trp Leu Asn Gly Lys Tyr Cys Lys Val
Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Ile Ser Lys
Ala Lys Gln Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Thr Ser
Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
Asn Gly Gln Met Leu Asn Asp Tyr Lys Thr Thr Pro Pro Leu
Asp Ser Asp Gly Ser Phe Leu Tyr Ser Lys Leu Thr Val Asp
Lys Ser Arg Trp Gln Gly Asn Val Phe Ser Cys Ser Val Met
His Gln Ala Leu His Asn His Tyr Thr Gin Lys Ser Leu Ser Leu
Ser Pro Gly Lys
(2) INFORMATION FOR SEQ ID NO:23:
(i) SEQUENCE CHARACTERISTICS:
   (3) LENGTH: 469 amino acids
   (4) TYPE: Amino Acid
(ii) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Gly Thr Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr
   1      5     10     15
Gly Val His Ser Glu Val Glu Leu Val Glu Ser Gly Gly Gly Leu
   20     25    30
Val Glu Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly
   35     40    45
Tyr Thr Phe Thr Glu Tyr Thr Met His Thr Met Thr Met Arg Glu Ala Pro
   50     55    60
Gly Lys Gly Leu Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly
   65     70    75
Gly Thr Ser His Asn Glu Arg Phe Met Asp Arg Phe Thr Ile Ser
   80     85    90
Val Asp Lys Ser Thr Ser Thr Tyr Met Glu Met Asn Ser Leu
   90    100    105
Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Thr Arg Gly
  110    115   120
Leu Asn Tyr Gly Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gin
  125    130   135
Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
  140    145   150
Val Phe Pro Leu Ala Pro Cys Ser Arg Ser Thr Ser Glu Ser Thr
  155    160    165
Ala Ala Leu Gly Cys Leu Val Asp Tyr Phe Pro Gln Pro Val
  170    175    180
Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr
  185    190    195
Phe Pro Ala Val Leu Glu Ser Gly Leu Tyr Ser Leu Ser Ser
  200    205    210
Val Val Thr Val Thr Ser Ser Asn Phe Gly Thr Gin Thr Tyr Thr
  215    220    225
Lys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys Thr
  230    235    240
Val Glu Arg Lys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro
  245    250    255
Phe Pro Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
  260    265    270
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
  275    280    285
Val Asp Val Ser His Glu Asp Pro Glu Val Glu Phe Asn Trp Tyr
  290    295    300
Val Asp Gly Met Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
  300    310    315
Glu Gln Phe Asn Thr Phe Arg Val Val Ser Val Leu Thr Val
  320    325    330
Val His Gln Asp Trp Leu Asn Gly Lys Tyr Lys Cys Lys Val
  335    340    345
<table>
<thead>
<tr>
<th>Ser</th>
<th>Asn</th>
<th>Lys</th>
<th>Gly</th>
<th>Leu</th>
<th>Pro</th>
<th>Ala</th>
<th>Pro</th>
<th>Ile</th>
<th>Gly</th>
<th>Lys</th>
<th>Thr</th>
<th>Ile</th>
<th>Ser</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350</td>
</tr>
<tr>
<td>Thr</td>
<td>Lys</td>
<td>Gly</td>
<td>Pro</td>
<td>Arg</td>
<td>Glu</td>
<td>Pro</td>
<td>Glu</td>
<td>Val</td>
<td>Tyr</td>
<td>Thr</td>
<td>Leu</td>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>365</td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Glu</td>
<td>Glu</td>
<td>Met</td>
<td>Thr</td>
<td>Lys</td>
<td>Asn</td>
<td>Glu</td>
<td>Val</td>
<td>Ser</td>
<td>Leu</td>
<td>Thr</td>
<td>Cys</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>380</td>
</tr>
<tr>
<td>Val</td>
<td>Lys</td>
<td>Gly</td>
<td>Phe</td>
<td>Tyr</td>
<td>Pro</td>
<td>Ser</td>
<td>Asp</td>
<td>Ile</td>
<td>Ala</td>
<td>Glu</td>
<td>Trp</td>
<td>Glu</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>395</td>
</tr>
<tr>
<td>Asn</td>
<td>Gly</td>
<td>Glu</td>
<td>Pro</td>
<td>Glu</td>
<td>Asn</td>
<td>Tyr</td>
<td>Lys</td>
<td>Thr</td>
<td>Thr</td>
<td>Pro</td>
<td>Pro</td>
<td>Met</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>410</td>
</tr>
<tr>
<td>Asp</td>
<td>Ser</td>
<td>Asp</td>
<td>Gly</td>
<td>Ser</td>
<td>Phe</td>
<td>Leu</td>
<td>Tyr</td>
<td>Ser</td>
<td>Leu</td>
<td>Thr</td>
<td>Val</td>
<td>Thr</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>425</td>
</tr>
<tr>
<td>Lys</td>
<td>Ser</td>
<td>Arg</td>
<td>Trp</td>
<td>Gin</td>
<td>Gly</td>
<td>Asn</td>
<td>Val</td>
<td>Phe</td>
<td>Ser</td>
<td>Ser</td>
<td>Cys</td>
<td>Ser</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>440</td>
</tr>
<tr>
<td>His</td>
<td>Glu</td>
<td>Ala</td>
<td>Leu</td>
<td>His</td>
<td>Asn</td>
<td>His</td>
<td>Tyr</td>
<td>Thr</td>
<td>Gin</td>
<td>Lys</td>
<td>Ser</td>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>455</td>
</tr>
<tr>
<td>Ser</td>
<td>Pro</td>
<td>Gly</td>
<td>Lys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>469</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:24:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 214 amino acids
      (B) TYPE: Amino Acid
      (D) TOPOLOGY: Linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

 Asp  Val  Gin  Met  Thr  Gin  Thr  Thr  Ser  Leu  Ser  Ala  Ser  Leu
 1   5  10  15
 Gly  Asp  Arg  Val  Thr  Asn  Cys  Arg  Ala  Ser  Gin  Asp  Ile  Asn
 20  25  30
 Asn  Tyr  Leu  Asn  Thr  Tyr  Gin  Glu  Gin  Lys  Pro  Asn  Gly  Thr  Val  Lys
 35  40  45
 Leu  Leu  Ile  Tyr  Tyr  Thr  Ser  Thr  Thr  Ser  Thr  Ser  Val  Pro  Ser
 50  55  60
 Arg  Phe  Ser  Gly  Ser  Gly  Thr  Asp  Tyr  Ser  Leu  Thr  Ile
 65  70  75
 Ser  Asn  Leu  Asp  Gin  Glu  Asp  Ile  Ala  Thr  Tyr  Phe  Ser  Asp  Ser
 80  85  90
 Gin  Thr  Thr  Leu  Pro  Pro  Pro  Thr  Gin  Gin  Gin  Thr  Leu  Tyr  Val  Glu
 95 100 105
 Ile  Lys  Arg  Thr  Val  Ala  Ala  Phe  Ser  Val  Glu  Phe  Phe  Pro  Pro
 110 115 120
 Ser  Asp  Glu  Glu  Ser  Gly  Thr  Ala  Ser  Val  Val  Cys  Leu
 125 130 135
 Leu  Asn  Asn  Phe  Thr  Asp  Pro  Arg  Glu  Ala  Lys  Val  Gin  Glu  Thr  Lys  Val
 140 145 150
 Asp  Asn  Ala  Leu  Glu  Ser  Gly  Asn  Ser  Gin  Glu  Ser  Val  Thr  Glu
 155 160 165
 Gin  Asp  Ser  Lys  Asp  Ser  Thr  Tyr  Ser  Leu  Ser  Ser  Thr  Leu  Thr
 170 175 180
 Leu  Ser  Lys  Ala  Asp  Tyr  Glu  Lys  His  Lys  Val  Tyr  Ala  Cys  Glu
 195 200 205
 Val  Thr  His  Gin  Gly  Leu  Ser  Ser  Ser  Pro  Val  Thr  Lys  Ser  Phe  Asn
 210 215 220
Arg Gly Glu Cys

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 233 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr
1   5   10   15
Gly Val His Ser Asp Ile Gin Met Thr Gin Ser Pro Ser Ser Ser Leu
20  25  30
Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser
35  40  45
Gln Asp Ile Asn Tyr Leu Asn Trp Tyr Gin Gin Lys Pro Gly
50  55  60
Lys Ala Pro Lys Leu Ile Tyr Tyr Thr Ser Thr Leu His Ser
65  70  75
Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr
80  85  90
Thr Leu Thr Ile Ser Ser Leu Gin Pro Gin Phe Thr Ala Thr
95 100 105
Tyr Cys Gin Gin Gly Asn Thr Leu Pro Pro Thr Phe Gin Gin
110 115 120
Thr Lys Val Glu Ile Lys Arg Thr Val Ala Pro Ser Val Phe
125 130 135
Ile Phe Pro Pro Ser Asp Glu Leu Lys Ser Gly Thr Ala Ser
140 145 150
Val Val Cys Leu Asn Phe Tyr Pro Arg Glu Ala Lys Val
155 160 165
Gln Trp Lys Val Asp Asn Leu Gin Ser Gin Ser Gin Gin
170 175 180
Ser Val Thr Glu Gin Ser Ser Asp Ser Thr Ser Ser Ser Leu Ser
185 190 195
Ser Thr Leu Thr Ser Lys Ala Asp Tyr Glu His Lys Val
200 205 210
Tyr Ala Cys Gin Val Thr His Gin Gly Leu Ser Ser Pro Val Thr
215 220 225
Lys Ser Phe Asp Arg Gly Glu Cys
230 233

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Glu Val Gin Leu Val Gin Ser Gly Gin Leu Val Gin Pro Gly
1   5   10   15
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr His Thr
20  25  30
Gln Tyr Thr Met His Thr Thr Thr Arg Gin Leu Ser Pro Gly Lys Gin

---continued---

35 40 45

Glu Trp Val Ala Gly Ile Asn Phe Lys Asn Gly Gly Thr Ser Tyr
50 55 60

Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Val Asp Lys Ser
65 70 75

Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
80 85 90

Thr Ala Val Tyr Tyr Cys Ala Arg Thr Arg Gly Leu Asn Tyr Gly
95 100 105

Phe Asp Val Asp Tyr Phe Asp Val Trp Gly Glu Gly Thr Leu Val
110 115 120

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
125 130 135

Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Gly
140 145 150

Cys Leu Val Lys Asp Tyr Phe Pro Gly Pro Val Thr Val Ser Trp
155 160 165

Asn Ser Gly Ala Leu Thr Ser Ser Gly Val His Thr Phe Pro Ala Val
170 175 180

Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
185 190 195

Pro Ser Ser Leu Gly Thr Gin Thr Tyr Ile Cys Asn Val Asn
200 205 210

His Lys Pro Ser Ser Thr Lys Val Asp Lys Lys Val Glu Pro Lys
215 220 225

Ser Cys Asp Lys Thr His Thr
230 235

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gin Pro Gly
1 5 10 15

Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20 25 30

Glu Tyr Thr Met His Thr Met Arg Gin Ala Pro Gly Lys Gly Leu
35 40 45

Glu Thr Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His
50 55 60

Asn Gin Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
65 70 75

Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
80 85 90

Thr Ala Val Tyr Cys Ala Arg Thr Arg Gly Leu Asn Tyr Gly
95 100 105

Phe Asp Val Asp Tyr Phe Asp Val Trp Gly Glu Gly Thr Leu Val
110 115 120

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
125 130 135

Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Gly
140 145 150

Cys Leu Val Lys Asp Tyr Phe Pro Gly Pro Val Thr Val Ser Trp
155 160 165

Asn Ser Gly Ala Leu Thr Ser Ser Gly Val His Thr Phe Pro Ala Val
170 175 180

Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
185 190 195

Pro Ser Ser Leu Gly Thr Gin Thr Tyr Ile Cys Asn Val Asn
200 205 210

His Lys Pro Ser Ser Thr Lys Val Asp Lys Lys Val Glu Pro Lys
215 220 225

Ser Cys Asp Lys Thr His Thr
230 235

---continued---
6,054,297

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(C) TOPOLOGY: Linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:28:
Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gin Pro Gly
1  5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Pro Thr
20  25 30
Glu Tyr Thr Met His Thr Met Arg Gin Ala Ala Pro Gly Lys Gly Leu
35  40 45
Glu Thr Val Ala Gly Ile Asp Pro Lys Gin Gly Gly Thr Ser His
50  55 60
Asn Gin Arg Pro Ser Arg Pro Pro Thr Leu Ala Val Asp Lys Ser
65  70 75
Lys Asn Thr Leu Tyr Leu Gin Met Asn Ser Leu Asp Ala Glu Asp
80  85 90
Thr Ala Val Tyr Cys Ala Arg Trp Arg Gly Leu Asp Tyr Gly
95 100 105
Phe Asp Val Arg Tyr Pro Val Gly Gin Gly Thr Leu Val
110 115 120
Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
125 130 135
Asp Pro Ser Ser Lys Ser Ser Gly Thr Ala Leu Gly
140 145 150
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
155 160 165
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
170 175 180
Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Thr Val
185 190 195
Pro Ser Ser Ser Leu Gly Thr Gin Thr Tyr Ile Cys Asn Val Asn
200 205 210
His Lys Pro Ser Ser Thr Lys Val Asp Lys Lys Val Glu Pro Lys
215 220 225
Ser Cys Asp Lys Thr His Thr
230 232
(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Glu Val Gin Leu Val Gin Ser Gly Gly Leu Val Gin Pro Gly
1  5  10
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20  25  30
Glu Tyr Thr Met His Thr Trp Met Arg Gin Ala Pro Gly Lys Gly Lys
35  40  45
Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His
50  55  60
Asn Gin Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
65  70  75
Thr Ser Thr Ala Tyr Met Gin Met Asn Ser Leu Arg Ala Glu Asp
80  85  90
Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
95 105
Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val
110 115 120
Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
125 130 135
Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
140 145 150
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
155 160 165
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
170 175 180
Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Val Val
185 190 195
Pro Ser Ser Ser Leu Gly Thr Gin Thr Tyr Ile Cys Asn Val Asn
200 205 210
His Lys Pro Ser Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
215 220 225
Ser Cys Asp Lys Thr His Thr
230 232

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 235 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Glu Val Gin Leu Val Gin Ser Gly Gly Leu Val Gin Pro Gly
1  5  10
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20  25  30
Glu Tyr Thr Met His Thr Trp Met Arg Gin Ala Pro Gly Lys Gly Lys
35  40  45
Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His
50  55  60
Asn Gin Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
65
70
75

Thr Ser Thr Ala Tyr Met Gin Met Asn Ser Leu Arg Ala Glu Asp
90
80
85
90

Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
95
100
100
105

Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val
110
115
120

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Ser Val Phe Pro Leu
125
130
135

Asa Pro Ser Ser Ser Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
140
145
150

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
150
160
165

Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
170
175
180

Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
185
190
195

Asp Ser Ser Ser Leu Gln Trp Thr Gin Thr Tyr Thr Cys Ser Val Asn
200
205
210

His Lys Pro Ser Ser Lys Leu Thr Asp Thr Lys Lys Val Glu Pro Lys
215
220
225

Ser Cys Asp Lys Thr His The Cys Pro Phe
230
235

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:31:
Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gin Pro Gly
1
5
10
15

Val Ser Leu Arg Leu Ser Cys Ala Thr Ser Uly Tyr Thr Phe Thr
20
25
30

Glu Tyr Thr Met His Thr Met Arg Gin Ala Pro Gly Leu Gly Leu
35
40
45

Glu Thr Val Ala Cys Ile Asn Pro Lys Asn Gly Gly Thr Ser His
50
55
60

Asn Gin Arg Phe Met Arg Arg Phe Thr Ile Ser Val Asp Lys Ser
65
70
75

Lys Asn Thr Leu Tyr Met Gin Leu Arg Ser Leu Arg Ala Glu Asp
80
85
90

Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
95
100
100
105

Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val
110
115
120

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Ser Val Phe Pro Leu
125
130
135

Asa Pro Ser Ser Lys Ser Ser Gly Gly Gly Thr Ala Ala Leu Gly
140
145
150

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
155
160
165
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gly Pro Gly
1 5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20 25 30
Glu Tyr Thr Met His Trp Met Arg Glu Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ala Gly Ile Asn Pro Lys Gly Gly Thr Ser His
50 55 60
Asn Gin Leu Thr Leu Leu Thr Ser Leu Thr Ser Glu Asp
65 70 75
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Thr Ser Glu Asp
80 85 90
Ser Gly Ile Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
95 100 105
Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val
110 115 120
Thr Val Ser Ser Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
125 130 135
Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
140 145 150
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
155 160 165
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Lys Pro Ala Val
170 175 180
Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
185 190 195
Pro Ser Ser Ser Leu Gly Thr Gin Thr Tyr Ile Cys Asn Val Asn
200 205 210
His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
215 220 225
Ser Cys Asp Lys Thr His Thr
230 232

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(C) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gly Pro Gly
1 5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20 25 30
Glu Tyr Thr Met His Trp Met Arg Glu Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ala Gly Ile Asn Pro Lys Gly Gly Thr Ser His
50 55 60
Asn Gin Leu Thr Leu Leu Thr Ser Leu Thr Ser Glu Asp
65 70 75
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Thr Ser Glu Asp
80 85 90
Ser Gly Ile Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
95 100 105
Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val
110 115 120
Thr Val Ser Ser Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
125 130 135
Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
140 145 150
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
155 160 165
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Lys Pro Ala Val
170 175 180
Leu Gin Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
185 190 195
Pro Ser Ser Ser Leu Gly Thr Gin Thr Tyr Ile Cys Asn Val Asn
200 205 210
His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
215 220 225
Ser Cys Asp Lys Thr His Thr
230 232

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gin Pro Gly
1  5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20 25 30
Glu Tyr Thr Met His Trp Met Arg Gin Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His
50 55 60
Asn Gin Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
65 70 75
Lys Asn Thr Leu Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80 85 90
Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
95 100 105
Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Ala Gly Thr Thr Val
110 115 120
Thr Val Ser Ser Ala Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
125 130 135
Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
140 145 150
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
155 160 165
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
170 175 180
Leu Gin Ser Ser Ile Leu Tyr Ser Leu Ser Ser Val Val Thr Val
185 190 195
Pro Ser Ser Ser Leu Lys Val Thr Thr Tyr Ile Cys Arg Val Asn
200 205 210
His Lys Pro Ser Ser Thr Leu Val Asp Lys Val Glu Pro Lys
215 220 225
Ser Cys Asp Lys Thr His Thr
230 232

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gin Pro Gly
1  5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
20 25 30
Glu Tyr Thr Met His Trp Met Arg Gin Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His
50 55 60
Asn Gin Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
65 70 75
Lys Asn Thr Leu Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80 85 90
Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly  
95  100  105

Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Glu Gly Thr Leu Val  
110  115  120

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu  
125  130  135

Asp Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly  
140  145  150

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp  
155  160  165

Asn Ser Gly Ala Leu Thr Ser Ser Gly Val His Thr Phe Pro Ala Val  
170  175  180

Leu Glu Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val  
185  190  195

Pro Ser Ser Leu Gly Thr Glu Thr Tyr Ile Cys Asn Val Asn  
200  205  210

His Lys Pro Ser Ser Thr Lys Val Asp Lys Lys Val Glu Pro Lys  
215  220  225

Ser Cys Asp Lys Thr His Thr  
230  232

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Glu Val Gln Leu Val Glu Ser Gly Gly Leu Val Lys Pro Gly  
1  5  10  15

Asn Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr  
20  25  30

Glu Tyr Thr Met His Trp Met Arg Glu Ala Pro Gly Lys Gly Leu  
35  40  45

Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His  
50  55  60

Asn Glu Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser  
65  70  75

Lys Asn Thr Leu Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp  
80  85  90

Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly  
95  100  105

Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Glu Gly Thr Leu Val  
110  115  120

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu  
125  130  135

Asp Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly  
140  145  150

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp  
155  160  165

Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val  
170  175  180

Leu Glu Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val  
185  190  195
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
Glu Tyr Thr Met His Trp Met Lys Gln Ser His Gly Lys Ser Leu
Glu Trp Val Ala Gly Ile Thr Pro Lys Asn Gly Gly Thr Ser His
Asn Glu Arg Phe Met Asn Phe Thr Ile Ser Val Asp Lys Ser
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gly Thr Leu Val
Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
 Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Thr Val Ser Trp
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
Leu Glu Ser Gly Leu Tyr Ser Ser Val Ser Val Thr Val
Pro Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn
Hist Lys Pro Ser Ser Lys Val Asp Lys Leu Glu Pro Lys
Ser Cys Asp Lys Thr His Thr

(2) INFORMATION FOR SEQ ID NO: 38:

(a) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

Glu Val Gln Leu Val Glu Ser Gly Gly Leu Val Gln Pro Gly
Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
Glu Tyr Thr Met His Trp Met Arg Gln Ala Pro Gly Lys Gly Leu
Glu Trp Ile Gly Gly Phe Asn Pro Lys Asn Gly Gly Thr Ser His
Asn Glu Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gly Thr Leu Val
Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
 Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Thr Val Ser Trp
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
Leu Glu Ser Gly Leu Tyr Ser Ser Val Ser Val Thr Val
Pro Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn
Hist Lys Pro Ser Ser Lys Val Asp Lys Leu Glu Pro Lys
Ser Cys Asp Lys Thr His Thr
6,054,297

-continued

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
Ala Pro Ser Ser Lys Ser Ser Thr Ser Gly Thr Ala Ala Leu Gly
Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Thr
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
Leu Gin Ser Ser Gly Leu Ser Ser Val Ser Val Thr Val
Pro Ser Ser Leu Gly Thr Gin Thr Ile Cys Asn Val Asn
His Lys Pro Ser Asn Thr Val Asp Lys Lys Val Glu Pro Lys
Ser Cys Asp Lys Thr His Thr

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 amino acids
(D) TYPE: Amino Acid
(T) TOPOLOGY: Linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:39:
Asp Ile Gin Met Thr Gin Ser Pro Ser Ser Leu Ser Ala Ser Val
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gin Asp Ile Asn
20 21 22 23 24 25 26 27 28 29 30
Asn Tyr Leu Asn Tyr Tyr Gin Gin Lys Pro Gly Lys Ala Pro Lys
35 36 37 38 39 40 41 42 43 44 45
Leu Leu Ile Tyr Thr Thr Leu Ser Ser Gin Val Pro Ser
50 51 52 53 54 55 56 57 58 59 60
Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile
65 66 67 68 69 70 71 72 73 74 75
Ser Ser Leu Gin Pro Glu Phe Ala Thr Tyr Tyr Cys Gin Gin
80 81 82 83 84 85 86 87 88 89 90
Gly Asn Thr Leu Pro Pro Thr Phe Gly Gin Gly Thr Lys Val Glu
95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115
Ile Lys Arg Thr Val Ala Asp Ser Val Pro Ser Val Phe Ile Phe Pro
116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
Ser Asp Glu Gin Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
Arg Gly Glu Cys
(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 amino acids
(B) TYPE: Amino Acid
(C) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Asp Ile Gin Met Thr Gin Ser Pro Ser Ser Leu Ser Ala Ser Val
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gin Asp Ile Asn
Asn Tyr Leu Asn Thr Tyr Gin Gin Lys Pro Gly Lys Ala Pro Lys
Leu Leu Ile Tyr Thr Ser Thr Leu His Ser Gly Val Pro Ser
Arg Phe Ser Ser Gin Ser Gin Thr Asp Tyr Thr Leu Thr Ile
Ser Ser Leu Gin Pro Glu Asp Phe Ala Thr Tyr Cys Gin Gin
Gly Asn Thr Leu Pro Pro Thr Phe Gin Gin Thr Lys Val Glu
Ile Lys Arg Thr Val Ala Pro Ser Val Phe Ile Phe Pro Pro
Ser Asp Gin Leu Lys Ser Gin Thr Ala Ser Val Val Cys Leu
Leu Asn Asn Asn Tyr Thr Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys
Asp Gin Ser Gin Ser Gin Ser Gin Ser Gin Ser Thr Thr
Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr
Leu Ser Lys Ala Asp Tyr Gin Lys Ala Asp Tyr Gin Lys Ala Cys Glu
Val Thr His Gin Leu Ser Ser Pro Val Thr Lys Ser Phe Asn
Arg Gly Gin Cys

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: Amino Acid
(C) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Asp Ile Gin Met Thr Gin Ser Pro Ser Ser Leu Ser Ala Ser Val
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gin Asp Ile Asn
Thr Ala Val Ala Trp Tyr Gin Gin Lys Pro Gly Lys Ala Pro Lys
Leu Leu Ile Tyr Ser Ala Ser Phe Thr Tyr Gin Gin Thr Lys Val Pro Ser
Arg Phe Ser Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile
65 70 75
Ser Ser Leu Gin Pro Gin Asp Phe Ala Thr Tyr Thr Cys Gin Gin
80 85 90
His Tyr Thr Thr Pro Thr Phe Gin Gin Thr Lys Val Glu
95 100 105
Ile Lys
107

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: Amino Acid
(D) TERTIARY TOPOLOGY: Linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Glu Val Gin Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly
1  5 10 15
Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Asn Ile Lys
20 25 30
Asp Thr Tyr Ile His Trp Val Arg Gin Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr
50 55 60
Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser
65 70 75
Lys Asp Thr Ala Tyr Leu Gin Met Ser Leu Arg Ala Gin Asp
80 85 90
Thr Ala Val Tyr Cys Ser Arg Trp Gly Gin Phe Val Tyr
95 100 105
Ala Met Asp Tyr Trp Gin Gly Thr Leu Val Thr Val Ser Ser
110 115 120

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: Amino Acid
(D) TERTIARY TOPOLOGY: Linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Asp Ile Gin Met Thr Gin Ser Ser Ser Ser Ala Ser Val
1  5 10 15
Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gin Asp Val Asn
20 25 30
Thr Ala Val Ala Trp Tyr Gin Gin Lys Pro Gin Gin Gin Gin Gin
35 40 45
Leu Leu Ile Tyr Ser Ala Ser Leu Met Ser Gin Val Pro Ser
50 55 60
Arg Phe Ser Gin Ser Gin Gin Thr Arg Thr Leu Thr Ile
65 70 75
Ser Ser Leu Gin Pro Gin Asp Phe Ala Thr Tyr Cys Gin Gin
80 85 90
His Tyr Thr Thr Pro Thr Phe Gin Gin Thr Lys Val Glu
95 100 105
Ile Lys
107
(2) INFORMATION FOR SEQ ID NO:44:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:
Glu Val Gin Leu Val Glu Ser Gly Gly Leu Val Gin Pro Gly
1  5  10  15
Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Asn Ile Lys
20  25  30
Asp Thr Tyr Ile His Trp Val Arg Gin Ala Pro Gly Lys Gly Leu
35  40  45
Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr
50  55  60
Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Xaa Asp Xaa Ser
65  70  75
Lys Asn Thr Xaa Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80  85  90
Thr Ala Val Tyr Cys Xaa Arg Trp Gly Gln Gly Phe Tyr
90  95  100
Ala Met Asp Xaa Trp Gly Gin Gly Thr Leu Val Thr Val Ser Ser
110 115 120

(2) INFORMATION FOR SEQ ID NO:45:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:
Glu Val Gin Leu Val Glu Ser Gly Gly Leu Val Gin Pro Gly
1  5  10  15
Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Ser Phe Thr
20  25  30
Gly Tyr Thr Met Asn Trp Val Arg Gin Ala Pro Gly Lys Gly Leu
35  40  45
Glu Trp Val Ala Leu Ile Asn Pro Tyr Lys Gly Val Thr Thr Tyr
50  55  60
Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Val Asp Lys Ser
65  70  75
Lys Asn Thr Ala Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80  85  90
Thr Ala Val Tyr Cys Ala Arg Ser Gly Tyr Cys Gly Arg Ser
95 100 105
Asp Trp Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val Thr Val
110 115 120

(2) INFORMATION FOR SEQ ID NO:46:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:
Glu Val Gin Leu Val Glu Ser Gly Gly Leu Val Gin Pro Gly
1  5  10  15
Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Ser Phe Thr
20  25  30
Gly Tyr Thr Met Asn Trp Val Arg Gin Ala Pro Gly Lys Gly Leu
35  40  45
Glu Trp Val Ala Leu Ile Asn Pro Tyr Lys Gly Val Thr Thr Tyr
50  55  60
Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Val Asp Lys Ser
65  70  75
Lys Asn Thr Ala Tyr Leu Gin Met Asn Ser Leu Arg Ala Glu Asp
80  85  90
Thr Ala Val Tyr Cys Ala Arg Ser Gly Tyr Cys Gly Arg Ser
95 100 105
Asp Trp Tyr Phe Asp Val Trp Gly Gin Gly Thr Leu Val Thr Val
110 115 120

Sex Ser
122
We claim:

1. A method for making a humanized antibody comprising non-human, import Complementarity Determining Region (CDR) amino acid residues and human Framework Region (FR) amino acid residues, comprising the steps of:
   (a) obtaining the amino acid sequences of an import variable domain and of a $\gamma_\mu$ subgroup III consensus human variable domain;
   (b) identifying CDR amino acid sequences in the import and the human variable domain sequences;
   (c) substituting import CDRs for the corresponding human CDRs;
   (d) aligning the amino acid sequences of a FR of the import antibody and the corresponding FR of the consensus variable domain;
   (e) identifying import antibody FR residues in the aligned FR sequences that are non-homologous to the corresponding consensus variable domain residues;
   (f) determining if the non-homologous import amino acid residue is expected to have at least one of the following effects:
      (1) non-covalently binds antigen directly;
      (2) interacts with a CDR; or
      (3) participates in the $\gamma_\mu$ $\gamma_\mu$ interface;
(g) for any non-homologous import antibody amino acid residue which is expected to have at least one of these effects, substituting that residue for the corresponding amino acid residue in the consensus variable domain FR sequence; and

(h) preparing a humanized antibody which binds antigen, wherein the humanized antibody comprises an amino acid sequence determined according to the above steps.

2. The method of claim 1, having an additional step of determining if any such non-homologous residues are exposed on the surface of the domain or buried within it, and if the residue is exposed, retaining the consensus residue.

3. The method of claim 1, having the additional steps of searching the import variable domain sequence for glycosylation sites, determining if any such glycosylation site is expected to affect the antigen binding or affinity of the antibody, and if so, substituting the glycosylation site into the consensus sequence.

4. The method of claim 1, having the additional steps of searching the consensus variable domain sequence for glycosylation sites which are not present at the corresponding amino acid in the import sequence, and if the glycosylation site is not present in the import sequence, substituting the import amino acid residues for the amino acid residues comprising the consensus glycosylation site.

5. The method of claim 1, having an additional step which comprises aligning import antibody and consensus variable domain FR sequences, identifying import antibody FR residues which are non-homologous with the aligned consensus FR sequence, and for each such non-homologous import antibody FR residue, determining if the corresponding consensus variable domain residue represents a residue which is highly conserved across all species at that site, and if it is so conserved, preparing a humanized antibody which comprises the consensus amino acid residue at that site.


7. A method for making a humanized antibody comprising non-human Complementarity Determining Region (CDR) amino acid residues and human Framework Region (FR) amino acid residues, comprising providing an import, non-human antibody variable domain amino acid sequence having CDR amino acid residues and FR amino acid residues; obtaining the amino acid sequence of a $\gamma_\text{H}$ subgroup III consensus human antibody variable domain having CDR amino acid residues and FR amino acid residues; substituting non-human CDR amino acid residues for human CDR amino acid residues in the consensus human antibody variable domain; substituting an amino acid residue for the consensus amino acid residue at least one of the following sites:


8. The method of claim 7, wherein the substituted residue is the residue found at the corresponding location of the non-human antibody.


10. The humanized antibody variable domain of claim 9, wherein the substituted residue is the residue found at the corresponding location of the non-human antibody from which the non-human CDR amino acid residues are obtained.

11. The humanized antibody variable domain of claim 9, wherein no human (FR) Framework Region (FR) residue other than those set forth in the group has been substituted.

12. A method for making a humanized antibody comprising introducing Complementarity Determining Region (CDR) amino acid residues from an import antibody variable domain into a $\gamma_\text{H}$ subgroup III consensus human antibody variable domain.

13. A humanized antibody variable domain having a functional antigen binding region, said humanized antibody variable domain comprising non-human Complementarity Determining Region (CDR) amino acid residues incorporated into a $\gamma_\text{H}$ subgroup III consensus human antibody variable domain and further comprising a non-human import Framework Region (FR) residue, wherein the non-human import FR residue introduces a glycosylation site which affects the antigen binding or affinity of the humanized antibody variable domain.

14. A humanized antibody which binds the HER2 receptor with an affinity of about $4.7 \text{ nM Kd}$ or better affinity and comprises a heavy chain variable domain which comprises non-human import antibody Complementarity Determining Region (CDR) amino acid residues incorporated into a $\gamma_\text{H}$ subgroup III consensus human variable domain.

15. A humanized variant of a non-human parent antibody, wherein the humanized variant comprises non-human Complementarity Determining Region (CDR) amino acid residues and human Framework Region (FR) amino acid residues, and:

(a) binds the HER2 receptor with an affinity of about $4.7 \text{ nM Kd}$ or better affinity;

(b) mediates specific cell lysis of SK-BR-3 cells in the presence of IL-2 activated human peripheral blood lymphocytes at least about four fold more effectively than the non-human parent antibody; and

(c) mediates Antibody Dependent Cellular Cytotoxicity (ADCC) selective for cell types which overexpress p185HER2 at least about two fold more effectively than for cell types which express low levels of p185HER2.

16. The humanized variant of claim 15 which binds the HER2 receptor with an affinity of about $0.82 \text{ nM Kd}$ or better affinity.

17. The humanized variant of claim 16 which binds the HER2 receptor with an affinity of about $0.10 \text{ nM Kd}$.

18. The humanized variant of claim 15 which inhibits proliferation of SK-BR-3 cells incubated for 96 hr with the antibody.

19. The humanized variant of claim 18 wherein the antibody inhibits proliferation of SK-BR-3 cells to about 66% of untreated control or greater inhibition.
20. A humanized variant of a non-human parent antibody, wherein the humanized variant comprises non-human Complementarity Determining Region (CDR) amino acid residues and human Framework Region (FR) amino acid residues, and;
(a) binds the HER2 receptor with an affinity of about 4.7 nM \( K_d \) or better affinity; and
(b) comprises a Framework Region (FR) amino acid substitution at a site selected from the group consisting of 71H, 73H, 78H, 93H and 66L, utilizing the numbering system set forth in Kabat.
21. The humanized variant of claim 20 which consists of about 1 to about 5 FR substitutions.
22. The humanized variant of claim 20 which comprises a FR substitution at site 73H.
23. The humanized variant of claim 20 which comprises a FR substitution at site 78H.
24. The humanized variant of claim 20 which comprises a FR substitution at site 93H.
25. The humanized variant of claim 20 which comprises a FR substitution at site 66L.
26. The humanized variant of claim 20 which further comprises a FR substitution at site 71H.
27. The humanized variant of claim 26 which comprises FR substitutions at sites 71H, 73H, 78H, 93H and 66L.
28. A humanized variant of a non-human parent antibody, wherein the humanized variant comprises non-human Complementarity Determining Region (CDR) amino acid residues and human Framework Region (FR) amino acid residues; binds the HER2 receptor with better affinity than the non-human parent antibody; and comprises Framework Region (FR) amino acid substitutions at sites 71H, 73H, 78H, 93H and 66L, utilizing the numbering system set forth in Kabat.
29. An antibody which binds an antigen and comprises a FR substitution at site 73H. Is said antigen and V\(_{\text{H}}\) subgroup III consensus human variable domain Framework Region (FR) amino acid residues; and comprises Framework Region (FR) amino acid residues; binds the HER2 receptor with better affinity than the non-human parent antibody; and comprises Framework Region (FR) amino acid substitutions at sites 71H, 73H, 78H, 93H and 66L, utilizing the numbering system set forth in Kabat.
30. The antibody of claim 29, further comprising V\(_{\text{L}}\) subgroup I consensus human variable domain FR amino acid residues.