List of peptides in the HIV-1 protease data sets

Index	Peptide	Cleavage	In 362 data set	In 746 data set	In 1625 data set	Number of Peptides	References
1	TQIMFETF	1	√	√	√	1	Poorman *et al.* (1991)
2	GQVNYEEF	1	√	√	√	1	Poorman *et al.* (1991)
3	PFIFEEEP	1	√	√	√	1	Poorman *et al.* (1991)
4	SFNFPQIT	1	√	√	√	1	Poorman *et al.* (1991)
5	DTVLEEMS	1	√	√	√	1	Poorman *et al.* (1991)
6	ARVLAEAM	1	√	√	√	1	Poorman *et al.* (1991)
7	AEELAEIF	1	√	√	√	1	Poorman *et al.* (1991)
8	SLNLRETN	1	√	√	√	1	Poorman *et al.* (1991)
9	ATIMMQRG	1	√	√	√	1	Poorman *et al.* (1991)
10	AECFRIFD	1	√	√	√	1	Poorman *et al.* (1991)
11	DQILIEIC	1	√	√	√	1	Poorman *et al.* (1991)
12	DDLFFEAD	1	√	√	√	1	Poorman *et al.* (1991)
13	YEEFVQMM	1	√	√	√	1	Poorman *et al.* (1991)
14	PIVGAETF	1	√	√	√	1	Poorman *et al.* (1991)
15	TLNFPISP	1	√	√	√	1	Poorman *et al.* (1991)
16	REAFRVFD	1	√	√	√	1	Poorman *et al.* (1991)
17	AETFYVDK	1	√	√		1	Poorman *et al.* (1991)
18	AQTFYVNL	1	√	√	√	1	Poorman *et al.* (1991)
19	PTLLTEAP	1	√	√	√	1	Poorman *et al.* (1991)
20	SFIGMESA	1	√	√	√	1	Poorman *et al.* (1991)
21	DAINTEFK	1	√	√	√	1	Poorman *et al.* (1991)
22	QITLWQRP	1	√	√	√	1	Poorman *et al.* (1991)
23	ELEFPEGG	1	√	√	√	1	Poorman *et al.* (1991)
24	SQNYPIVQ	1	√	√	√	1	Poorman *et al.* (1991)
25	PGNFLQSR	1	√	√	√	1	Poorman *et al.* (1991)
26	GDALLERN	1	√	√	√	1	Poorman *et al.* (1991)
27	KELYPLTS	1	√	√	√	1	Poorman *et al.* (1991)
28	RQANFLGK	1	√	√	√	1	Poorman *et al.* (1991)
29	SRSLYASS	1	√	√	√	1	Poorman *et al.* (1991)
30	AEAMSQVT	1	√	√	√	1	Poorman *et al.* (1991)
31	RKILFLDG	1	√	√	√	1	Poorman *et al.* (1991)
32	GSHLVEAL	1	√	√	√	1	Poorman *et al.* (1991)
33	GGVYATRS	1	√	√	√	1	Poorman *et al.* (1991)
34	FRSGVETT	1	√	√	√	1	Poorman *et al.* (1991)
35	VEVAEEEE	1	√	√	√	1	Poorman *et al.* (1991)
36	LPVNGEFS	1	√	√	√	1	Poorman *et al.* (1991)
37	ETTALVCD	1	√	√	√	1	Poorman *et al.* (1991)
38	HLVEALYL	1	√	√	√	1	Poorman *et al.* (1991)
39	HYGFPTYG	1	√	√	√	1	Tomasselli *et al.* (1993)
40	DSADAEED	1	√	√	√	1	Tomasselli *et al.* (1993)
41	GWILGEHG	1	√	√	√	1	Tomasselli *et al.* (1993)
42	GWILAEHG	1	√	√	√	1	Tomasselli *et al.* (1993)
43	QAIYLALQ	1	√	√	√	1	Tomasselli *et al.* (1993)
44	EKVYLAWV	1	√	√	√	1	Tomasselli *et al.* (1993)
45	VEICTEME	1	√	√	√	1	Tomasselli *et al.* (1993)
46	TQDFWEVQ	1	√	√	√	1	Tomasselli *et al.* (1993)
47	LWMGYELH	1	√	√	√	1	Tomasselli *et al.* (1993)
48	GDAYFSVP	1	√	√	√	1	Tomasselli *et al.* (1993)
49	ELELAENR	1	√	√	√	1	Tomasselli *et al.* (1993)
50	SKDLIAEI	1	√	√	√	1	Tomasselli *et al.* (1993)
51	LEVNIVTD	1	√	√	√	1	Tomasselli *et al.* (1993)
52	GGNYPVQH	1	√	√	√	1	Tomasselli *et al.* (1993)
53	ARLMAEAL	1	√	√	√	1	Tomasselli *et al.* (1993)
54	PFAAAQQR	1	√	√	√	1	Tomasselli *et al.* (1993)
55	PRNFPVAQ	1	√	√	√	1	Tomasselli *et al.* (1993)
56	GLAAPQFS	1	√	√	√	1	Tomasselli *et al.* (1993)
57	SLNLPVAK	1	√	√	√	1	Tomasselli *et al.* (1993)

58	AETFYTDG	1	√	√	√	1	Tomasselli *et al.* (1993)
59	RQVLFLEK	1	√	√	√	1	Tomasselli *et al.* (1993)
60	QMIFEEHG	1	√	√	√	1	Tomasselli *et al.* (1993)
61	RQNYPIVQ	1	√	√	√	1	Tözsér *et al.* (1991b)
62	SQNYAIVQ	1	√	√	√	1	Partin *et al.* (1990)
63	TQNYPIVQ	1	√	√	√	1	Partin *et al.* (1990)
64	SNNYPIVQ	1	√	√		1	Partin *et al.* (1990)
65	SKNYPIVQ	1	√	√		1	Partin *et al.* (1990)
66	SQNFPIVQ	1	√	√	√	1	Partin *et al.* (1990)
67	SQNYLIVQ	1	√	√		1	Partin *et al.* (1990)
68	SQNYTIVQ	1	√	√		1	Partin *et al.* (1990)
69	SQNYPIIQ	1	√	√		1	Partin *et al.* (1990)
70	SQNYPIEQ	1	√	√		1	Partin *et al.* (1990)
71	SQNYPIVP	1	√	√		1	Partin *et al.* (1990)
72	SQNYPIVE	1	√	√		1	Partin *et al.* (1990)
73	TFNFPQIT	1	√	√		1	Partin *et al.* (1990)
74	YFNFPQIT	1	√	√		1	Partin *et al.* (1990)
75	SCNFPQIT	1	√	√		1	Partin *et al.* (1990)
76	SYNFPQIT	1	√	√		1	Partin *et al.* (1990)
77	SFTFPQIT	1	√	√		1	Partin *et al.* (1990)
78	SFYFPQIT	1	√	√		1	Partin *et al.* (1990)
79	SFNSPQIT	1	√	√		1	Partin *et al.* (1990)
80	SFNYPQIT	1	√	√		1	Partin *et al.* (1990)
81	SFNFGQIT	1	√	√		1	Partin *et al.* (1990)
82	SFNFLQIT	1	√	√		1	Partin *et al.* (1990)
83	SFNFPPIT	1	√	√		1	Partin *et al.* (1990)
84	SFNFPLIT	1	√	√		1	Partin *et al.* (1990)
85	SFNFPQVT	1	√	√		1	Partin *et al.* (1990)
86	SFNFPQDT	1	√	√		1	Partin *et al.* (1990)
87	SFNFPQII	1	√	√		1	Partin *et al.* (1990)
88	SQNYPAVQ	1	√	√		1	Partin *et al.* (1990)
89	SQNYPLVQ	1	√	√	√	1	Partin *et al.* (1990)
90	SQNYPVVQ	1	√	√		1	Partin *et al.* (1990)
91	SQNYPNVQ	1	√	√		1	Thompson *et al.* (1995)
92	SQNYPILQ	1	√	√	√	1	Cai *et al.* (1998)
93	SQNYPIFQ	1	√	√	√	1	Cai *et al.* (1998)
94	SQLYPIVQ	1	√	√	√	1	Cai *et al.* (1998)
95	SQCYPIVQ	1	√	√	√	1	Cai *et al.* (1998)
96	SQAYPIVQ	1	√	√	√	1	Cai *et al.* (1998)
97	SQTYPIVQ	1	√	√	√	1	Cai *et al.* (1998)
98	SQNMPIVQ	1	√	√	√	1	Cai *et al.* (1998)
99	ARVLFEAL	1	√	√		1	Chou (1993)
100	ARVLFQAL	1	√	√		1	Chou (1993)
101	ARVLFIAL	1	√	√		1	Chou (1993)
102	ARVLFVAL	1	√	√		1	Chou (1993)
103	ARVLFAAL	1	√	√		1	Chou (1993)
104	ARVLFDAL	1	√	√		1	Chou (1993)
105	ARVLFNAL	1	√	√		1	Chou (1993)
106	ARVLFTAL	1	√	√		1	Chou (1993)
107	ARNLFEAL	1	√	√		1	Chou (1993)
108	ARNLFQAL	1	√	√		1	Chou (1993)
109	ARNLFIAL	1	√	√		1	Chou (1993)
110	ARNLFVAL	1	√	√		1	Chou (1993)
111	ARVYPEAL	1	√	√		1	Chou (1993)
112	ARNYPEAL	1	√	√		1	Chou (1993)
113	SQNYPIVL	1	√	√		1	Chou (1993)
114	ARNYPIVL	1	√	√		1	Chou (1993)
115	AQNYPIVL	1	√	√		1	Chou (1993)
116	RQNYPIAL	1	√	√		1	Chou (1993)
117	SQNYDIVQ	0	√	√	√	1	Chou (1993)
118	SQNYKIVQ	0	√	√	√	1	Chou (1993)
119	SQKYPIVQ	0	√	√	√	1	Partin *et al.* (1990)
120	SQQYPIVQ	0	√	√		1	Partin *et al.* (1990)

121	SQNSPIVQ	0	√	√	√	1	Partin *et al.* (1990)
122	SQNPPIVQ	0	√	√	√	1	Partin *et al.* (1990)
123	SQNYPKVQ	0	√	√	√	1	Partin *et al.* (1990)
124	P00704	0	√	√	√	122	Chou *et al.* (1996)
125	P61824	0	√	√	√	117	Chou *et al.* (1996)
126	AAVLAEAM	1		√	√	1	Ridky *et al.* (1998)
127	AFVLAEAM	1		√	√	1	Ridky *et al.* (1998)
128	AGVLAEAM	1		√	√	1	Ridky *et al.* (1998)
129	AKVLAEAM	1		√	√	1	Ridky *et al.* (1998)
130	AQVLAEAM	1		√	√	1	Ridky *et al.* (1998)
131	ARALAEAM	1		√	√	1	Ridky *et al.* (1998)
132	ARGLAEAM	1		√	√	1	Ridky *et al.* (1998)
133	ARILAEAM	1		√	√	1	Ridky *et al.* (1998)
134	ARRLAEAM	1		√	√	1	Ridky *et al.* (1998)
135	ARVAAEAM	1		√	√	1	Ridky *et al.* (1998)
136	ARVFAEAM	1		√	√	1	Ridky *et al.* (1998)
137	ARVIAEAM	1		√	√	1	Ridky *et al.* (1998)
138	ARVLAEAA	1		√	√	1	Ridky *et al.* (1998)
139	ARVLAEAE	1		√	√	1	Ridky *et al.* (1998)
140	ARVLAEAG	1		√	√	1	Ridky *et al.* (1998)
141	ARVLAEAK	1		√	√	1	Ridky *et al.* (1998)
142	ARVLAEAP	1		√	√	1	Ridky *et al.* (1998)
143	ARVLAEAR	1		√	√	1	Ridky *et al.* (1998)
144	ARVLAEEM	1		√	√	1	Ridky *et al.* (1998)
145	ARVLAEFM	1		√	√	1	Ridky *et al.* (1998)
146	ARVLAEGM	1		√	√	1	Ridky *et al.* (1998)
147	ARVLAERM	1		√	√	1	Ridky *et al.* (1998)
148	ARVLAESM	1		√	√	1	Ridky *et al.* (1998)
149	ARVLALAM	1		√	√	1	Ridky *et al.* (1998)
150	ARVLAVAM	1		√	√	1	Ridky *et al.* (1998)
151	ARVLFEAM	1		√	√	1	Ridky *et al.* (1998)
152	ARVLGEAM	1		√	√	1	Ridky *et al.* (1998)
153	ARVLIEAM	1		√	√	1	Ridky *et al.* (1998)
154	ARVLLEAM	1		√	√	1	Ridky *et al.* (1998)
155	ARVLMEAM	1		√	√	1	Ridky *et al.* (1998)
156	ARVLNEAM	1		√	√	1	Ridky *et al.* (1998)
157	ARVLVEAM	1		√	√	1	Ridky *et al.* (1998)
158	ARVMAEAM	1		√	√	1	Ridky *et al.* (1998)
159	ARVNAEAM	1		√	√	1	Ridky *et al.* (1998)
160	ARVYAEAM	1		√	√	1	Ridky *et al.* (1998)
161	ATVLAEAM	1		√	√	1	Ridky *et al.* (1998)
162	ERVLAEAM	1		√	√	1	Ridky *et al.* (1998)
163	MRVLAEAM	1		√	√	1	Ridky *et al.* (1998)
164	PRVLAEAM	1		√	√	1	Ridky *et al.* (1998)
165	RRVLAEAM	1		√	√	1	Ridky *et al.* (1998)
166	ARVLAQAM	0		√	√	1	Ridky *et al.* (1998)
167	ARVVAEAM	0		√		1	Ridky *et al.* (1998)
168	ARNLAEAM	0		√		1	Ridky *et al.* (1998)
169	ARSLAEAM	0		√		1	Ridky *et al.* (1998)
170	ARVLAAAM	0		√		1	Ridky *et al.* (1998)
171	ARVLADAM	0		√		1	Ridky *et al.* (1998)
172	ARVLAEDM	0		√		1	Ridky *et al.* (1998)
173	ARVLAGAM	0		√		1	Ridky *et al.* (1998)
174	ARVLANAM	0		√		1	Ridky *et al.* (1998)
175	GAVWLAMT	1		√	√	1	Ridky *et al.* (1996)
176	PAASFAMT	1		√	√	1	Ridky *et al.* (1996)
177	PAASLAMT	1		√	√	1	Ridky *et al.* (1996)
178	PAAWLAMT	1		√	√	1	Ridky *et al.* (1996)
179	PAGSFAMT	1		√	√	1	Ridky *et al.* (1996)
180	PAGWLAMT	1		√	√	1	Ridky *et al.* (1996)
181	PAHSFAMT	1		√	√	1	Ridky *et al.* (1996)
182	PAHSLAMT	1		√	√	1	Ridky *et al.* (1996)
183	PALSFAMT	1		√	√	1	Ridky *et al.* (1996)

184	PALWLAMT	1	√	√	1	Ridky *et al.* (1996)
185	PAVIFAMT	1	√	√	1	Ridky *et al.* (1996)
186	PAVILAMT	1	√	√	1	Ridky *et al.* (1996)
187	PAVLAAMT	1	√	√	1	Ridky *et al.* (1996)
188	PAVLFAMT	1	√	√	1	Ridky *et al.* (1996)
189	PAVLGAMT	1	√	√	1	Ridky *et al.* (1996)
190	PAVSAAMT	1	√	√	1	Ridky *et al.* (1996)
191	PAVSVAMT	1	√	√	1	Ridky *et al.* (1996)
192	PAVVLAMT	1	√	√	1	Ridky *et al.* (1996)
193	PAVWAAMT	1	√	√	1	Ridky *et al.* (1996)
194	PAVWFAMT	1	√	√	1	Ridky *et al.* (1996)
195	PAVWGAMT	1	√	√	1	Ridky *et al.* (1996)
196	PAVWVAMT	1	√	√	1	Ridky *et al.* (1996)
197	PGVWLAMT	1	√	√	1	Ridky *et al.* (1996)
198	PHVWLAMT	1	√	√	1	Ridky *et al.* (1996)
199	PGLSLAMT	1	√	√	1	Ridky *et al.* (1996)
200	PRASLAMT	1	√	√	1	Ridky *et al.* (1996)
201	PRGSLAMT	1	√	√	1	Ridky *et al.* (1996)
202	PRHSLAMT	1	√	√	1	Ridky *et al.* (1996)
203	PRLSLAMT	1	√	√	1	Ridky *et al.* (1996)
204	PRVALAMT	1	√	√	1	Ridky *et al.* (1996)
205	PRVGLAMT	1	√	√	1	Ridky *et al.* (1996)
206	PRVLLAMT	1	√	√	1	Ridky *et al.* (1996)
207	PRVWLAMT	1	√	√	1	Ridky *et al.* (1996)
208	PSVSLAMT	1	√	√	1	Ridky *et al.* (1996)
209	PSVWLAMT	1	√	√	1	Ridky *et al.* (1996)
210	PTVSLAMT	1	√	√	1	Ridky *et al.* (1996)
211	PTVWLAMT	1	√	√	1	Ridky *et al.* (1996)
212	PASWLAMT	1	√		1	Ridky *et al.* (1996)
213	FAVSLAMT	1	√	√	1	Beck *et al.* (2002)
214	GAVSLAMT	1	√	√	1	Beck *et al.* (2002)
215	HAVSLAMT	1	√	√	1	Beck *et al.* (2002)
216	LAVSLAMT	1	√	√	1	Beck *et al.* (2002)
217	NAVSLAMT	1	√	√	1	Beck *et al.* (2002)
218	PAGSLAMT	1	√	√	1	Beck *et al.* (2002)
219	PASSLAMT	1	√	√	1	Beck *et al.* (2002)
220	PALSLAMT	1	√	√	1	Beck *et al.* (2002)
221	PAVALAMT	1	√	√	1	Beck *et al.* (2002)
222	PAVELAMT	1	√	√	1	Beck *et al.* (2002)
223	PAVGLAMT	1	√	√	1	Beck *et al.* (2002)
224	PAVHLAMT	1	√	√	1	Beck *et al.* (2002)
225	PAVLLAMT	1	√	√	1	Beck *et al.* (2002)
226	PAVRLAMT	1	√	√	1	Beck *et al.* (2002)
227	PAVSEAMT	1	√	√	1	Beck *et al.* (2002)
228	PAVSFAMT	1	√	√	1	Beck *et al.* (2002)
229	PAVSGAMT	1	√	√	1	Beck *et al.* (2002)
230	PAVSLALT	1	√	√	1	Beck *et al.* (2002)
231	PAVSLAMT	1	√	√	1	Beck *et al.* (2002)
232	PAVSLANT	1	√	√	1	Beck *et al.* (2002)
233	PAVSLAYT	1	√	√	1	Beck *et al.* (2002)
234	PAVSLGMT	1	√	√	1	Beck *et al.* (2002)
235	PAVSLHMT	1	√	√	1	Beck *et al.* (2002)
236	PAVSLLMT	1	√	√	1	Beck *et al.* (2002)
237	PAVSLSMT	1	√	√	1	Beck *et al.* (2002)
238	PAVSQAMT	1	√	√	1	Beck *et al.* (2002)
239	PAVSRAMT	1	√	√	1	Beck *et al.* (2002)
240	PAVWLAMT	1	√	√	1	Beck *et al.* (2002)
241	PAWSLAMT	1	√	√	1	Beck *et al.* (2002)
242	PDVSLAMT	1	√	√	1	Beck *et al.* (2002)
243	PFVSLAMT	1	√	√	1	Beck *et al.* (2002)
244	PGVSLAMT	1	√	√	1	Beck *et al.* (2002)
245	PHVSLAMT	1	√	√	1	Beck *et al.* (2002)
246	PNVSLAMT	1	√	√	1	Beck *et al.* (2002)

247	PPVSLAMT	1	√	√	1	Beck *et al.* (2002)
248	PRVSLAMT	1	√	√	1	Beck *et al.* (2002)
249	GQANFLGK	1	√	√	1	Feher *et al.* (2002)
250	GQVNFLGK	1	√	√	1	Feher *et al.* (2002)
251	PGNFFQSR	1	√	√	1	Feher *et al.* (2002)
252	PGNFPQSR	1	√	√	1	Feher *et al.* (2002)
253	PGNFVQSR	1	√	√	1	Feher *et al.* (2002)
254	PGNLLQSR	1	√	√	1	Feher *et al.* (2002)
255	PRNFLQSR	1	√	√	1	Feher *et al.* (2002)
256	PGNYLQSR	1	√	√	1	Feher *et al.* (2002)
257	RQVNFLGK	1	√	√	1	Feher *et al.* (2002)
258	RRANFLGK	1	√	√	1	Feher *et al.* (2002)
259	QGNFLQSR	1	√	√	1	Feher *et al.* (2002)
260	RRVNFLGK	1	√	√	1	Feher *et al.* (2002)
261	AQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
262	DQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
263	KQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
264	LQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
265	MQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
266	NQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
267	GQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
268	PQNYPIVQ	1	√	√	1	Tözsér *et al.* (1991b)
269	DKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
270	GKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
271	LKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
272	AKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
273	PKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
274	SKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
275	SQIYPIVQ	1	√	√	1	Tözsér *et al.* (2000)
276	SQVYPIVQ	1	√	√	1	Tözsér *et al.* (2000)
277	TAVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
278	TDVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
279	TFVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
280	TGVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
281	TKALVVQP	1	√	√	1	Tözsér *et al.* (2000)
282	TKDLVVQP	1	√	√	1	Tözsér *et al.* (2000)
283	TKFLVVQP	1	√	√	1	Tözsér *et al.* (2000)
284	TKILVVQP	1	√	√	1	Tözsér *et al.* (2000)
285	TKLLVVQP	1	√	√	1	Tözsér *et al.* (2000)
286	TKNLVVQP	1	√	√	1	Tözsér *et al.* (2000)
287	TKVAVVQP	1	√	√	1	Tözsér *et al.* (2000)
288	TKVFVVQP	1	√	√	1	Tözsér *et al.* (2000)
289	TKVGVVQP	1	√	√	1	Tözsér *et al.* (2000)
290	TKVLFVQP	1	√	√	1	Tözsér *et al.* (2000)
291	TKVLIVQP	1	√	√	1	Tözsér *et al.* (2000)
292	TKVLLVQP	1	√	√	1	Tözsér *et al.* (2000)
293	TKVLPVQP	1	√	√	1	Tözsér *et al.* (2000)
294	TKVLSVQP	1	√	√	1	Tözsér *et al.* (2000)
295	TKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
296	TKVMVVQP	1	√	√	1	Tözsér *et al.* (2000)
297	TKVYVVQP	1	√	√	1	Tözsér *et al.* (2000)
298	TLVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
299	TSVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
300	TVVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
301	VKVLVVQP	1	√	√	1	Tözsér *et al.* (2000)
302	KTKKLVVQPK	0	√		3	Tözsér *et al.* (2000)
303	TKKLVVQP	0		√	1	Tözsér *et al.* (2000)
304	KTKVKVVQPK	0	√		3	Tözsér *et al.* (2000)
305	TKVKVVQP	0		√	1	Tözsér *et al.* (2000)
306	KTKVLKVQPK	0	√		3	Tözsér *et al.* (2000)
307	TKVLKVQP	0		√	1	Tözsér *et al.* (2000)
308	ATAMMATA	1	√	√	1	Tözsér *et al.* (1997)
309	ATIMMITA	1	√	√	1	Tözsér *et al.* (1997)

310	ATIYYITA	1	√	√	1	Tözsér *et al.* (1997)
311	SDAYYADS	1	√	√	1	Tözsér *et al.* (1997)
312	SDAYYTDS	1	√	√	1	Tözsér *et al.* (1997)
313	SDCYYCDS	1	√	√	1	Tözsér *et al.* (1997)
314	SDCYYTDS	1	√	√	1	Tözsér *et al.* (1997)
315	SDEYYEDS	1	√	√	1	Tözsér *et al.* (1997)
316	SDEYYTDS	1	√	√	1	Tözsér *et al.* (1997)
317	SDGYYTDS	1	√	√	1	Tözsér *et al.* (1997)
318	SDIYYTDS	1	√	√	1	Tözsér *et al.* (1997)
319	SDLYYTDS	1	√	√	1	Tözsér *et al.* (1997)
320	SDNYPIVQ	1	√	√	1	Tözsér *et al.* (1997)
321	SDTYYADS	1	√	√	1	Tözsér *et al.* (1997)
322	SDTYYCDS	1	√	√	1	Tözsér *et al.* (1997)
323	SDTYYEDS	1	√	√	1	Tözsér *et al.* (1997)
324	SDTYYGDS	1	√	√	1	Tözsér *et al.* (1997)
325	SDTYYIDS	1	√	√	1	Tözsér *et al.* (1997)
326	SDTYYLDS	1	√	√	1	Tözsér *et al.* (1997)
327	SDTYYTDQ	1	√	√	1	Tözsér *et al.* (1997)
328	SDTYYTDS	1	√	√	1	Tözsér *et al.* (1997)
329	SDTYYTFS	1	√	√	1	Tözsér *et al.* (1997)
330	SDTYYTGS	1	√	√	1	Tözsér *et al.* (1997)
331	SDTYYTLS	1	√	√	1	Tözsér *et al.* (1997)
332	SDTYYTQS	1	√	√	1	Tözsér *et al.* (1997)
333	SFTYYTDS	1	√	√	1	Tözsér *et al.* (1997)
334	SGTYYTDS	1	√	√	1	Tözsér *et al.* (1997)
335	SGTYYTGS	1	√	√	1	Tözsér *et al.* (1997)
336	SLTYYTDS	1	√	√	1	Tözsér *et al.* (1997)
337	SQNYPIDQ	1	√	√	1	Tözsér *et al.* (1997)
338	SQNYPTVQ	1	√	√	1	Tözsér *et al.* (1997)
339	SQNYYIDQ	1	√	√	1	Tözsér *et al.* (1997)
340	SQNYYIVQ	1	√	√	1	Tözsér *et al.* (1997)
341	SQNYYNQS	1	√	√	1	Tözsér *et al.* (1997)
342	SQNYYTDQ	1	√	√	1	Tözsér *et al.* (1997)
343	SQNYYTVQ	1	√	√	1	Tözsér *et al.* (1997)
344	SQTYYIDQ	1	√	√	1	Tözsér *et al.* (1997)
345	SQTYYIVQ	1	√	√	1	Tözsér *et al.* (1997)
346	SQTYYTDQ	1	√	√	1	Tözsér *et al.* (1997)
347	SQTYYTDS	1	√	√	1	Tözsér *et al.* (1997)
348	SQTYYTQS	1	√	√	1	Tözsér *et al.* (1997)
349	SQTYYTVQ	1	√	√	1	Tözsér *et al.* (1997)
350	GEMFFPVL	1	√	√	1	Beck *et al.* (2000)
351	GFAMAEAL	1	√	√	1	Beck *et al.* (2000)
352	GFIGVSYL	1	√	√	1	Beck *et al.* (2000)
353	GLNMPALV	1	√	√	1	Beck *et al.* (2000)
354	GLTMVQEL	1	√	√	1	Beck *et al.* (2000)
355	GLVAFANL	1	√	√	1	Beck *et al.* (2000)
356	GLVLQEGL	1	√	√	1	Beck *et al.* (2000)
357	GQDFPMYL	1	√	√	1	Beck *et al.* (2000)
358	GWVMTEAL	1	√	√	1	Beck *et al.* (2000)
359	NSVMIALV	1	√	√	1	Beck *et al.* (2000)
360	SGAFMTRG	1	√	√	1	Beck *et al.* (2000)
361	SGAYLIQG	1	√	√	1	Beck *et al.* (2000)
362	SGIFLETS	1	√	√	1	Beck *et al.* (2000)
363	SGIMFESN	1	√	√	1	Beck *et al.* (2000)
364	SGIMFQSA	1	√	√	1	Beck *et al.* (2000)
365	SGINFESG	1	√	√	1	Beck *et al.* (2000)
366	SGIYLVEN	1	√	√	1	Beck *et al.* (2000)
367	SGIYYSVS	1	√	√	1	Beck *et al.* (2000)
368	SGLYYVTE	1	√	√	1	Beck *et al.* (2000)
369	SGMWFEAP	1	√	√	1	Beck *et al.* (2000)
370	SGNFAAFS	1	√	√	1	Beck *et al.* (2000)
371	SGNMLVYS	1	√	√	1	Beck *et al.* (2000)
372	SGNMVMFG	1	√	√	1	Beck *et al.* (2000)

373	SGNYFVET	1	√	√	1	Beck *et al.* (2000)
374	SGNYFVQG	1	√	√	1	Beck *et al.* (2000)
375	SGNYLVTS	1	√	√	1	Beck *et al.* (2000)
376	SGSYVEYQ	1	√	√	1	Beck *et al.* (2000)
377	SGTFQVQL	1	√	√	1	Beck *et al.* (2000)
378	SGVFTEER	1	√	√	1	Beck *et al.* (2000)
379	SGVFVEMP	1	√	√	1	Beck *et al.* (2000)
380	SGVFVETS	1	√	√	1	Beck *et al.* (2000)
381	SGVFVVNG	1	√	√	1	Beck *et al.* (2000)
382	SGVFYSRE	1	√	√	1	Beck *et al.* (2000)
383	SGVHFISR	1	√	√	1	Beck *et al.* (2000)
384	SGVHVEYT	1	√	√	1	Beck *et al.* (2000)
385	SGVLFVSS	1	√	√	1	Beck *et al.* (2000)
386	SGVMFQTD	1	√	√	1	Beck *et al.* (2000)
387	SGVMPTMS	1	√	√	1	Beck *et al.* (2000)
388	SGVYHVST	1	√	√	1	Beck *et al.* (2000)
389	SGVYLATD	1	√	√	1	Beck *et al.* (2000)
390	SFVFVVNG	1	√	√	1	Beck *et al.* (2001)
391	SGIFVVNG	1	√	√	1	Beck *et al.* (2001)
392	SGIYTVQS	1	√	√	1	Beck *et al.* (2001)
393	SGLFTEYG	1	√	√	1	Beck *et al.* (2001)
394	SGNFVVNG	1	√	√	1	Beck *et al.* (2001)
395	SGVFVENG	1	√	√	1	Beck *et al.* (2001)
396	SGVFVVQG	1	√	√	1	Beck *et al.* (2001)
397	SGVFVVTG	1	√	√	1	Beck *et al.* (2001)
398	SGVFAVTQ	1	√	√	1	Beck *et al.* (2001)
399	SGVFHVNG	1	√	√	1	Beck *et al.* (2001)
400	SGVFQVNG	1	√	√	1	Beck *et al.* (2001)
401	SGVFVQNG	1	√	√	1	Beck *et al.* (2001)
402	AETFYVDG	1	√	√	1	Beck *et al.* (2001)
403	FHVNGLVK	0	√		1	Beck *et al.* (2001)
404	FQVNGLVK	0	√		1	Beck *et al.* (2001)
405	KSGVFSVNGLVK	0	√		5	Beck *et al.* (2001)
406	SGVFSVNG	0		√	1	Beck *et al.* (2001)
407	KGSGVYQLSALVPK	0	√		7	Beck *et al.* (2001)
408	SGVYQLSA	0		√	1	Beck *et al.* (2001)
409	KGSGGRINVALVPK	0	√		7	Beck *et al.* (2001)
410	GRINVALV	0		√	1	Beck *et al.* (2001)
411	KSGVFVNNGLVK	0	√	√	5	Beck *et al.* (2001)
412	KGSGALTNAVLVPK	0	√	√	7	Beck *et al.* (2001)
413	KGSGAMVNQALVPK	0	√	√	7	Beck *et al.* (2001)
414	KGSGTWMVHSLVPK	0	√	√	7	Beck *et al.* (2001)
415	KGSGLTMVTQLVPK	0	√	√	7	Beck *et al.* (2001)
416	KSGVNVVNGK	0	√	√	3	Beck *et al.* (2001)
417	KSRVNVVNGK	0	√	√	3	Beck *et al.* (2001)
418	KSGVQVVNGK	0	√	√	3	Beck *et al.* (2001)
419	KSRVQVVNGK	0	√	√	3	Beck *et al.* (2001)
420	GVALSALV	0	√		1	Beck *et al.* (2001)
421	FVVNGLVK	0	√		1	Beck *et al.* (2001)
422	TLNFPQIT	1	√		1	Beck *et al.* (2001)
423	SGVFYTLV	0	√		1	Beck *et al.* (2001)
424	AAAMSSAI	0	√	√	1	Kadas *et al.* (2004)
425	SSLYPALT	1	√	√	1	Kadas *et al.* (2004)
426	TFTFPVVF	1	√	√	1	Kadas *et al.* (2004)
427	ATVLTVAL	1	√	√	1	Kadas *et al.* (2004)
428	DLVLLSAE	1	√	√	1	Kadas *et al.* (2004)
429	EEIMLAYQ	1	√	√	1	Kadas *et al.* (2004)
430	ELILPVKR	1	√	√	1	Kadas *et al.* (2004)
431	FQAYPLRE	1	√	√	1	Kadas *et al.* (2004)
432	KDIFPVTE	1	√	√	1	Kadas *et al.* (2004)
433	KLVLAQLS	1	√	√	1	Kadas *et al.* (2004)
434	KMMLLAKA	1	√	√	1	Kadas *et al.* (2004)
435	LECLLSIP	1	√	√	1	Kadas *et al.* (2004)

436	PAILPIIS	1	√	√	1	Kadas *et al.* (2004)
437	PLIMAVVN	1	√	√	1	Kadas *et al.* (2004)
438	PMVGVLDA	1	√	√	1	Kadas *et al.* (2004)
439	PQVLPVMH	1	√	√	1	Kadas *et al.* (2004)
440	PVILPIQA	1	√	√	1	Kadas *et al.* (2004)
441	SEEYPIMI	1	√	√	1	Kadas *et al.* (2004)
442	SQAFPLRA	1	√	√	1	Kadas *et al.* (2004)
443	SKAFLADT	1	√	√	1	Kadas *et al.* (2004)
444	LQVLTLNI	0	√	√	1	Kadas *et al.* (2004)
445	PAILVHTP	0	√	√	1	Kadas *et al.* (2004)
446	PYVGSGLY	0	√	√	1	Kadas *et al.* (2004)
447	SKLLATVV	0	√	√	1	Kadas *et al.* (2004)
448	STLLIENS	0	√	√	1	Kadas *et al.* (2004)
449	TSLLTLDD	0	√	√	1	Kadas *et al.* (2004)
450	VVAMPVVI	0	√	√	1	Kadas *et al.* (2004)
451	ASILPVIP	0	√	√	1	Kadas *et al.* (2004)
452	KQTFPIQQ	0	√	√	1	Kadas *et al.* (2004)
453	TSCYCHGT	1	√		1	Kadas *et al.* (2004)
454	QANFLGKI	0	√		1	Kurt *et al.* (2003)
455	QITLPKRP	1	√		1	Kurt *et al.* (2003)
456	RKVLFLDG	1	√		1	Kurt *et al.* (2003)
457	PQNFLQSR	1	√	√	1	Tözsér *et al.* (1991a)
458	RQAGFLGL	0		√	1	Tözsér *et al.* (1991a)
459	ATIMMQRE	1		√	1	Cameron *et al.* (1992)
460	PLFAGISE	1		√	1	Cameron *et al.* (1992)
461	PLIMANVN	1		√	1	Cameron *et al.* (1992)
462	SQNYPQVQ	1		√	1	Boross *et al.* (1999)
463	ARNYPIAQ	1		√	1	Boross *et al.* (1999)
464	SRNYPEVQ	1		√	1	Boross *et al.* (1999)
465	SRVLAEAM	1		√	1	Boross *et al.* (1999)
466	SRVLAIAM	1		√	1	Boross *et al.* (1999)
467	SQNYPEAQ	1		√	1	Boross *et al.* (1999)
468	ARVLAIAM	1		√	1	Boross *et al.* (1999)
469	SQNYPEVQ	1		√	1	Boross *et al.* (1999)
470	ARNYPQAQ	1		√	1	Boross *et al.* (1999)
471	SQNYPIAQ	1		√	1	Boross *et al.* (1999)
472	SRNYPEAQ	1		√	1	Boross *et al.* (1999)
473	ARNYPEAQ	1		√	1	Boross *et al.* (1999)
474	SRNYPIAQ	1		√	1	Boross *et al.* (1999)
475	SRNYPIVQ	1		√	1	Boross *et al.* (1999)
476	SQFYPIVQ	1		√	1	Tözsér *et al.* (1992)
477	SQNVPIVQ	0		√	1	Tözsér *et al.* (1992)
478	SQNIPIVQ	0		√	1	Tözsér *et al.* (1992)
479	SQNGPIVQ	0		√	1	Tözsér *et al.* (1992)
480	SQNDPIVQ	0		√	1	Tözsér *et al.* (1992)
481	SQNKPIVQ	0		√	1	Tözsér *et al.* (1992)
482	SDGYYGDS	0		√	1	Tözsér *et al.* (1997)
483	SQNPYIVQ	0		√	1	Tözsér *et al.* (1997)
484	QVIPYNQS	0		√	1	Tözsér *et al.* (1997)
485	PAHWLAMT	1		√	1	Ridky *et al.* (1996)
486	LFAGISDW	0		√	1	Tözsér *et al.* (1996)
487	TSCYHCGT	1		√	1	Kadas *et al.* (2004)
488	P08670(1-52)	0		√	45	Kontijevskis *et al.* (2007)
489	P08670(53-61)	0		√	2	Kontijevskis *et al.* (2007)
490	P08670(62-93)	0		√	25	Kontijevskis *et al.* (2007)
491	P08670(94-423)	0		√	323	Kontijevskis *et al.* (2007)
492	P08670(424-466)	0		√	36	Kontijevskis *et al.* (2007)
493	RT sequence	0		√	494	Kontijevskis *et al.* (2007)
494	PIHDHDHPFHGYQLEKEP	0		√	11	Kontijevskis *et al.* (2007)

- An 8-mer denotes $P_4 - P_{4'}$ and the cleavage site is between the P_1 and $P_{1'}$. For peptides longer than eight residues and protein sequences, a sliding window with eight residue long is used to extract all possible 8-mers along them.

- All the I(Ile) and F(Phe) amino acids in the $99^{th} - 116^{th}$ entries were actually Nle and Nph in Griffiths *et al.* (1992), where Nph is the chromophoric reporter group, *p*-nitrophenylalanine and Nle is the Norleucine.
- The 124^{th} and 125^{th} entries are based on hen egg lysozyme c (with access number P00704 in SwissProt) and ribonuclease pancreatic (with access number P61824 in SwissProt).
- With the 302^{th}, 304^{th}, 306^{th}, 405^{th}, 407^{th} and 409^{th} entries, a sliding window was used on them to extract 8-mers in the 746 data set. However, only one 8-mer for each entry was extracted in the 1625 data set.
- The 487^{th} peptide should be TSCYCHGT according to the corresponding reference.
- The $488^{th} - 492^{th}$ entries are based on vimentin protein (with access number p08670 in SwissProt). The numbers inside the parentheses are the corresponding amino acid positions in the vimentin protein sequence.
- The 493^{th} entry is based on Reverse Transcriptase.
- The 494^{th} entry is part of RNase H protein.