**Supplementary materials**

Structural catalytic core of the members of the superfamily of acid proteases

Alexander I. Denesyuka,Konstantin Denessiouka, Mark S. Johnsona, Vladimir N. Uverskyb

a Structural Bioinformatics Laboratory, Biochemistry, InFLAMES Research Flagship Center, Faculty of Science and Engineering, Biochemistry, Åbo Akademi University, Turku 20520, Finland

b Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA

**Table S1**. Conserved geometric parameters (distance and angle) of contacts in 33 DD-zones of the acid proteases superfamily proteins.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| N | PDB ID | Protein  | DD-zone (Circular DD-links & D-loops connection) | Fireman’s grip |
| Superfamily: Acid proteases |
| Family: Pepsin-like |
| 1a, b | 3PSG\_A | Propepsin | O/Y14-HOH507 2.7 HOH507-N/F31 2.8 | OG/S36-O/Y125 2.7N/Y125-O/G188 3.2 | O/W190-N/V214 2.9 | OG/S219-O/D11 2.9 | O/F31-OG1/T216 2.7N/T33-OG1/T216 2.8OG1/T33-N/T216 2.9OG1/T33-O/V214 3.1 |
| 1c, d | 4PEP\_A | Pepsin  | O/Y14-HOH426 3.2HOH426-N/F31 3.1 | OG/S36-O/Y125 2.6N/Y125-O/G188 2.8 | O/W190-N/V214 2.9 | OG/S219-O/T12 3.0 | O/F31-OG1/T216 2.6N/T33-OG1/T216 3.0OG1/T33-N/T216 3.0OG1/T33-O/V214 2.9 |
| 1e, f | 6XCZ\_A | Pepsin  | O/Y14-N/F31 3.8 | OG/S36-O/Y125 2.7N/Y125-O/G188 2.9 | O/W190-N/V214 3.0 | OG/S219-O/T12 2.4 | O/F31-OG1/T216 2.6N/T33-OG1/T216 3.0OG1/T33-N/T216 2.9OG1/T33-O/V214 2.8 |
| 2a, b | 3VCM\_A | Prorenin  | O/Y14-N/F31 5.2 | OG/S36-O/F125 2.7N/F125-O/G188 4.2 | O/W190-N/V214 3.0 | CB/S219-O/D11 3.2 (2.5) 120O | O/F31-OG1/T216 2.6N/T33-OG1/T216 2.9OG1/T33-N/T216 2.9OG1/T33-O/V214 2.7 |
| 2c, d | 2REN\_A | Renin  | O/Y20-N/F37 4.0 | OG/S42-O/F132 3.2N/F132-O/G199 4.3 | O/W201-N/V225 3.0 | OG/S230-O/T18 3.3 | O/F37-OG1/T227 2.7N/T39-OG1/T227 2.8OG1/T39-N/T227 2.9OG1/T39-O/V225 2.6 |
| 2e, f | 3K1W\_A | Renin  | O/Y20-HOH357 2.8HOH357-N/F37 2.9 | OG/S42-O/F132 2.7N/F132-HOH387 3.0HOH387-O/G199 2.9 | O/W201-N/V225 2.9 | OG/S230-O/T18 2.7 | O/F37-OG1/T227 2.7N/T39-OG1/T227 2.9OG1/T39-N/T227 2.8OG1/T39-O/V225 2.7 |
| 3a, b | 1PFZ\_A | Proplasmepsin 2 | O/M15-N/L33 2.9 | CB/A38-O/W128 3.0 (2.0) 169ON/W128-HOH469 2.8HOH469-O/L191 2.6 | O/W193-N/V213 3.0 | OG/S218-O/Q12 3.1 | O/L33-OG/S215 6.4N/T35-OG/S215 6.0OG1/T35-N/S215 4.9OG1/T35-O/V213 4.5 |
| 3c, d | 1LF4\_A | Plasmepsin 2  | O/F16-HOH339 2.7HOH339-N/L33 2.8 | CB/A38-O/W128 3.3 (2.7) 113ON/W128-O/L191 3.7 | O/W193-N/V213 3.1 | CB/S218-O/I14 3.2 (2.6) 114O | O/L33-OG/S215 2.6N/T35-OG/S215 3.0OG1/T35-N/S215 2.9OG1/T35-O/V213 3.3 |
| 3e, f | 2BJU\_A | Plasmepsin 2  | O/F16-HOH2051 2.7HOH2051-N/L33 2.9 | CB/A38-O/W128 3.4 (3.0) 104ON/W128-O/L191 3.1 | O/W193-N/V213 2.9 | CB/S218-O/I14 2.9 (2.2) 115O | O/L33-OG/S215 2.6N/T35-OG/S215 2.9OG1/T35-N/S215 2.9OG1/T35-O/V213 2.9 |
| 4a, b | 3QVC\_A | HAP zymogen  | O/L13-N/F31 2.8 | OG/S36-O/W125 3.7N/W125-HOH398 2.9HOH398-O/L188 2.5 | O/W190-N/L214 2.8 | OG/S219-O/L9 2.9 | O/F31-OG/S216 7.7N/T33-OG/S216 7.3OG1/T33-N/S216 5.0OG1/T33-HOH430 2.9HOH430-O/L214 2.9 |
| 4e, f | 3QVI\_A | HAP protein  | O/S14-N/F31 2.9 | OG/S36-O/W125 2.6N/W125-HOH590 2.5HOH590-O/L188 2.6 | O/W190-N/L214 2.9 | OG/S219-CD1/L13 3.8 (3.2) 113OCD2/L13-O/V122.8 (1.8) 156O | O/F31-OG/S216 2.9N/T33-OG/S216 2.8OG1/T33-N/S216 2.9OG1/T33-O/L214 3.1 |
| 5a, b  | 5N7N\_A | Procathepsin D  | O/Y40-N/F57 4.1 | CB/A62-O/Y1513.3 (2.7) 113ON/Y151-O/G217 2.9 | O/W219-N/A248 3.1 | OG/S253-O/D37 4.1 | O/F57-OG1/T250 2.8N/T59-OG1/T250 2.9OG1/T59-N/T250 2.9OG1/T59-O/A248 2.9 |
| 5c, d | 5N71\_A | Cathepsin D  | O/Y40-HOH565 2.8 HOH565-N/F57 2.8 | CB/A62-O/Y1513.5 (2.8) 117ON/Y151-O/G217 3.0 | O/W219-N/A248 3.2 | OG/S253-O/V38 3.4 | O/F57-OG1/T250 2.6N/T59-OG1/T250 2.8OG1/T59-N/T250 2.9OG1/T59-O/A248 2.8 |
| 5e, f | 5N7Q\_A | Cathepsin D  | O/Y18-HOH595 2.7 HOH595-N/F35 2.8 | CB/A40-O/Y1293.4 (2.7) 124ON/Y129-O/G195 3.0 | O/W197-N/A226 3.0 | OG/S231-O/V16 2.6 | O/F35-OG1/T228 2.7N/T37-OG1/T228 2.9OG1/T37-N/T228 2.9OG1/T37-O/A226 2.8 |
| 6a, b | 1MIQ\_A | Proplasmepsin | O/M15-N/F33 3.1 | CB/A38-O/W128 3.5 (2.5) 147ON/W128-HOH362 2.7HOH362-O/L191 2.8 | O/W193-N/V213 3.1 | OG1/T218-ND2/N13 3.4  | O/F33-OG/S215 7.0N/T35-OG/S215 6.7OG1/T35-N/S215 5.1OG1/T35-HOH342 2.7HOH342-O/V213 3.0 |
| 6e, f | 1QS8\_A | Plasmepsin  | O/F16-HOH350 3.1 HOH350-N/F33 2.7 | CB/A38-O/W128 3.2 (2.7) 105ON/W128-O/L191 3.1 | O/W193-N/V213 2.9 | CG2/T218-O/I14 4.0 (3.4) 114O | O/F33-OG/S215 2.5N/T35-OG/S215 3.0OG1/T35-N/S215 3.0OG1/T35-O/V213 2.7 |
| 7a, b | 5JOD\_A | Proplasmepsin 4 | O/M15-N/F33 3.1 | CB/A38-O/W128 4.3 (3.6) 122ON/W128-HOH513 2.9HOH513-O/L191 2.5 | O/W193-N/V213 3.0 | CB/S218-OD1/N13 3.3 (2.7) 118O  | O/F33-OG/S215 7.4N/T35-OG/S215 6.9OG1/T35-N/S215 5.1OG1/T35-HOH681 3.0HOH681-O/V213 2.9 |
| 7e, f | 1LS5\_A | Plasmepsin 4  | O/F16-N/F33 4.8 | CB/A38-O/W128 3.6 (3.3) 98ON/W128-O/L191 3.0 | O/W193-N/V213 3.1 | CB/S218-O/L14 2.9 (2.1) 123O | O/F33-OG/S215 2.9N/T35-OG/S215 2.8OG1/T35-N/S215 3.2OG1/T35-O/V213 2.9 |
| 8a, b | 1QDM\_A | Prophytepsin | O/Y18-N/F35 4.9 | OG/S40-O/F130 2.5N/F130-O/G195 3.5 | O/W197-N/A222 3.1 | OG/S227-O/N15 2.8 | O/F35-OG/S224 3.6N/T37-OG/S224 3.4OG1/T37-N/S224 2.9OG1/T37-O/A222 3.0 |
| 9a, b | 1HTR\_B | Progastricsin | O/Y14-HOH378 2.7 HOH378-N/F31 2.8 | CB/S36-CB/Y125 4.9CA/Y125-O/L189 3.9 (2.8) 161O | O/W191-N/V216 2.9 | OG/S221-HOH396 3.1HOH396-O/D11 2.9 | O/F31-OG1/T218 2.7N/T33-OG1/T218 3.0OG1/T33-N/T218 2.8OG1/T33-O/V216 2.8 |
| 10a, b | 1TZS\_A | Procathepsin E | O/Y25-HOH388 2.6 HOH388-N/F42 2.9 | OG/S47-O/Y136 2.6N/Y136-O/A201 2.9 | O/W203-N/V227 3.1 | OG/S232-O/M23 2.6OG/S232-O/D22 3.3 | O/F42-OG1/T229 2.6N/T44-OG1/T229 3.0OG1/T44-N/T229 3.0OG1/T44-O/V227 2.8 |
| 11c, d | 1T6E\_X | Xylanase inhibitor | O/Y16-N/L28 2.8 | O/P33-OG/S146 2.7N/N145-O/P204 2.7 | O/H206-N/L234 3.0 | CG/P239-CB/S14 4.4 | O/L28-OG1/T236 2.8N/V30-OG1/T236 3.1CG2/V30-N/T236 3.6CG2/V30-O/L234 3.4 (2.6) 130O |
| 11e, f | 1T6G\_A | Xylanase inhibitor | O/Y16-N/L28 2.8 | O/P33-OG/S146 3.5N/N145-O/P204 2.9 | O/H206-N/L234 3.1 | CG/P239-CB/S14 4.3 | O/L28-OG1/T236 2.7N/V30-OG1/T236 3.0CG2/V30-N/T236 3.7CG2/V30-O/L234 3.4 (2.6) 133O |
| 12c, d | 3AUP\_A | Basic 7S globulin | O/H23-N/V40 2.8 | OD1/N45-CB/A164 3.2 (2.6) 112O N/H163-HOH512 3.0HOH512-O/G228 2.7 | O/Y230-N/I264 3.2 | CB/P269-O/G214.0 (3.0) 146O | O/V40-OG1/T266 2.6N/L42-OG1/T266 3.0CB/L42-N/T266 3.9CD2/L42-O/I2643.4 (2.3) 175O |
| 13c, d | 3VLA\_A | EDGP (Fragment) | O/Y22-N/V39 2.8 | O/R44-OG1/T160 2.7N/R159-O/V235 2.9 | O/Y237-N/I270 3.1 | CB/P275-O/L204.4 (3.4) 158O | O/V39-OG1/T272 2.8N/L41-OG1/T272 3.0CB/L41-N/T272 4.2CD2/L41-O/I2703.4 (2.4) 156O |
| 13e, f | 3VLB\_A | EDGP (Fragment) | O/Y22-N/V39 3.0 | O/R44-OG1/T160 2.8N/R159-O/V235 3.1 | O/Y237-N/I270 2.9 | CB/P275-O/L204.2 (3.2) 152O | O/V39-OG1/T272 2.5N/L41-OG1/T272 3.3CB/L41-N/T272 4.4CD2/L41-O/I2703.6 (2.6) 146O |
| Family: Retroviral protease (retropepsin) |
| 14c, d | 3IXO\_A, B | HIV-1 protease | O/P9\_A-N/L24\_A 2.8 | OD1/D29\_A-NH2/R87\_A 2.5OD1/D29\_A-NE/R8\_B 3.3OD2/D29\_A-NH2/R8\_B 3.3 | O/P9\_B-N/L24\_B 3.0 | OD1/D29\_B-NH2/R87\_B 3.5OD1/D29\_B-NE/R8\_A 3.2OD2/D29\_B-NH2/R8\_A 2.8 | O/L24\_A-OG1/T26\_B 2.9N/T26\_A-OG1/T26\_B 2.9OG1/T26\_A-N/T26\_B 2.9OG1/T26\_A-O/L24\_B 2.5 |
| 14e, f | 5YOK\_A, B | HIV-1 protease | O/P9\_A-N/L24\_A 2.9 | OD1/D29\_A-NH2/R87\_A 2.9OD1/D29\_A-NE/R8\_B 3.0OD2/D29\_A-NH2/R8\_B 2.8 | O/P9\_B-N/L24\_B 2.9 | OD1/D29\_B-NH2/R87\_B 2.9OD1/D29\_B-NE/R8\_A 2.8OD2/D29\_B-NH2/R8\_A 2.9 | O/L24\_A-OG1/T26\_B 2.7N/T26\_A-OG1/T26\_B 2.9OG1/T26\_A-N/T26\_B 2.9OG1/T26\_A-O/L24\_B 2.7 |
| 15c, d | 3NR6\_A, B | XMRV protease | O/P16\_A-N/V31\_A 2.9 | OE1/Q36\_A-NH2/R95\_A 5.0CD/R95\_A-O/E15\_B 3.1 (2.6) 108O | O/P16\_B-N/V31\_B 2.9 | OE1/Q36\_B-CB/R95\_B 3.9 (3.0) 150OCD/R95\_B-O/E15\_A 3.0 (2.6) 104O | O/V31\_A-OG1/T33\_B 2.6N/T33\_A-OG1/T33\_B 3.0OG1/T33\_A-N/T33\_B 2.9OG1/T33\_A-O/V31\_B 2.6 |
| 15e, f | 3SLZ\_A, B | XMRV protease | O/P16\_A-N/V31\_A 2.8 | OE1/Q36\_A-HOH178 3.0 HOH178-NH2/R95\_A 2.9CD/R95\_A-O/E15\_B 3.3 (2.8) 108O | O/P16\_B-N/V31\_B 2.9 | OE1/Q36\_B-NH2/R95\_B 4.6CD/R95\_B-O/E15\_A 3.2 (2.6) 114O | O/V31\_A-OG1/T33\_B 2.6N/T33\_A-OG1/T33\_B 2.9OG1/T33\_A-N/T33\_B 2.9OG1/T33\_A-O/V31\_B 2.7 |
| Family: Dimeric aspartyl proteases |
| 16c, d | 4Z2Z\_A, B | Ddi1 protease  | O/L204\_A-N/V219\_A 2.8 | NE2/Q224\_A-O/V201\_B 3.4 | O/L204\_B-N/V219\_B 2.8 | NE2/Q224\_B-O/V201\_A 4.2 | O/V219\_A-OG/T221\_B 3.0N/T221\_A-OG1/T221\_B 2.9OG1/T221\_A-N/T221\_B 2.8OG1/T221\_A-N/V219\_B 2.9 |
| 17c | 5C9F\_A | ApRick protease | O/F124-N/V139 2.9 | N/A | N/A | N/A | N/A |
| Family: LPG0085-like |
| 18c, d | 2PMA\_A, I | Protein Lpg0085 | O/Y29\_A-N/L46\_A 2.8 | NZ/K51\_A-OD1/D148\_A 3.0R147\_A-Y29\_I 3.9 (π-π) |  O/Y29\_I-N/L46\_I 2.7 | NZ/K51\_I-OD2/D148\_I 3.2R147\_I-Y29\_A 3.9 (π-π) | O/L46\_A-OG1/T48\_I 2.7N/T48\_A-OG1/T48\_I 2.9OG1/T48\_A-N/T48\_I 3.0OG1/T48\_A-O/L46\_I 2.7 |

N/A – Not Available.

**Table S2**. Conserved geometric parameters (distance and angle) of contacts in 65 psi-loops of the acid proteases superfamily proteins and contacts between DD-linkN and the propeptide/N-terminal peptide in 13 pepsin-like family proteins.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N | PDB ID | Protein  | psi-loop (D-loop & G-loop)  | Asx-motif | DD-linkN/Propep.DD-linkN/N-pep. |
| Superfamily: Acid proteases |
| Family: Pepsin-like |
| 1a | 3PSG\_A, p | Propepsin | N/D32-O/L121 3.1O/D32-N/L123 2.8O/S35-N/A124 2.7 | O/D32-CA/G122 3.5 (2.6) 143O | OD1/D32-N/G34 2.8O/D32-N/S35 3.3 | O/E13-N/K9P 2.8N/F15-O/V7P 3.0O/F15-N/V7P 2.7 |
| 1b | 3PSG\_A | Propepsin | N/D215-O/L301 3.0O/T218-N/D303 2.9O/S219-N/V304 3.1 | O/D215-CA/G302 3.4 (2.4) 157OO/G302-N/F305 3.0 | OD1/D215-N/G217 2.6O/D215-N/T218 3.4 |  |
| 1c | 4PEP\_A | Pepsin  | N/D32-O/L121 2.9O/D32-N/L123 2.9O/S35-N/A124 2.8 | O/D32-CA/G122 3.5 (2.6) 141O | OD1/D32-N/G34 3.1O/D32-N/S35 3.4 | O/E13-N/Y9 2.8N/F15-O/E7 3.2O/F15-N/E7 2.7 |
| 1d | 4PEP\_A | Pepsin  | N/D215-O/L301 2.9O/T218-N/D303 2.7O/S219-N/V304 3.0 | O/D215-CA/G302 3.5 (2.4) 163OO/G302-N/F305 3.0 | OD2/D215-N/G217 3.0O/D215-N/T218 3.3 |  |
| 1e | 6XCZ\_A | Pepsin  | N/D32-O/L121 2.9O/D32-N/L123 2.9O/S35-N/A124 3.0 | O/D32-CA/G122 3.4 (2.5) 146O | OD1/D32-N/G34 3.0O/D32-N/S35 3.4 | O/E13-N/Y9 3.0N/F15-O/E7 3.3O/F15-N/E7 2.8 |
| 1f | 6XCZ\_A | Pepsin  | N/D215-O/L301 3.1O/T218-N/D303 2.8O/S219-N/V304 3.2 | O/D215-CA/G302 3.5 (2.4) 162OO/G302-N/F305 2.9 | OD1/D215-N/G217 2.9O/D215-N/T218 3.2 |  |
| 2a | 3VCM\_A, p | Prorenin  | N/D32-O/V121 3.1O/D32-N/M123 2.7O/S35-N/G124 2.8 | O/D32-CA/G122 3.5 (2.6) 139O | OD1/D32-N/G34 2.9O/D32-N/S35 3.7 | O/Q13-N/M16P 2.7N/Y15-O/K14P 3.1O/Y15-N/K14P 2.9 |
| 2b | 3VCM\_A | Prorenin  | N/D215-O/L301 3.2O/A218-N/A303 2.5O/S219-N/T304 3.2 | O/D215-CA/G302 3.3 (2.2) 170OO/G302-N/F305 3.1 | OD1/D215-N/G217 3.0O/D215-N/A218 3.5 |  |
| 2c | 2REN\_A | Renin  | N/D38-O/V128 3.1O/D38-N/M130 2.8O/S41-N/G131 2.9 | O/D38-CA/G129 3.4 (2.6) 125O | OD1/D38-N/G40 2.8O/D38-N/S41 3.0 | O/Q19-N/Y15 2.6N/Y21-O/T13 2.9O/Y21-N/T13 2.9 |
| 2d | 2REN\_A | Renin  | N/D226-O/L315 2.9O/A229-N/A317 2.5O/S230-N/T318 3.5 | O/D226-CA/G316 2.9 (2.1) 153OO/G316-N/F319 3.0 | OD1/D226-N/G228 3.3O/D226-N/A229 3.3 |  |
| 2e | 3K1W\_A | Renin  | N/D38-O/V128 2.9O/D38-N/M130 2.8O/S41-N/G131 2.8 | O/D38-CA/G129 3.5 (2.5) 141O | OD1/D38-N/G40 3.1O/D38-N/S41 3.4 | O/Q19-N/Y15 2.8N/Y21-O/T13 3.1O/Y21-N/T13 2.8 |
| 2f | 3K1W\_A | Renin  | N/D226-O/L315 2.9O/A229-N/A317 2.9O/S230-N/T318 3.4 | O/D226-CA/G316 3.3 (2.2) 155OO/G316-N/F319 2.9 | OD1/D226-N/G228 3.0O/D226-N/A229 3.1 |  |
| 3a | 1PFZ\_A, p | Proplasmepsin 2 | N/D34-O/L124 3.1O/D34-N/L126 2.9O/S37-N/G127 2.8 | O/D34-CA/G125 3.6 (2.6) 144O | OD1/D34-N/G36 2.7O/D34-N/S37 3.3 | O/I14-N/E87P 2.9N/F16-O/K85P 2.8O/F16-N/K85P 3.0 |
| 3b | 1PFZ\_A | Proplasmepsin 2 | N/D214-O/L301 3.1O/T217-N/D303 2.8O/S218-CD/P304 3.4 (2.6) 131O | O/D214-CA/G302 3.2 (2.2) 156OO/G302-N/F305 3.0 | OD1/D214-N/G216 2.8O/D214-N/T217 3.2 |  |
| 3c | 1LF4\_A | Plasmepsin 2  | N/D34-O/L124 3.0O/D34-N/L126 2.7O/S37-N/G127 2.8 | O/D34-CA/G125 3.4 (2.5) 139O | OD2/D34-N/G36 2.8O/D34-N/S37 3.3 | O/M15-N/F11 2.7N/Y17-O/V9 2.8O/Y17-N/V9 3.0 |
| 3d | 1LF4\_A | Plasmepsin 2  | N/D214-O/L301 3.1O/T217-N/D303 2.8O/S218-CD/P304 3.2 (2.3) 135O | O/D214-CA/G302 3.2 (2.2) 153OO/G302-N/F305 3.0 | OD1/D214-N/G216 2.8O/D214-N/T217 3.5 |  |
| 3e | 2BJU\_A | Plasmepsin 2  | N/D34-O/L124 2.9O/D34-N/L126 2.8O/S37-N/G127 2.8 | O/D34-CA/G125 3.4 (2.5) 142O | OD2/D34-N/G36 2.8O/D34-N/S37 3.3 | O/M15-N/F11 2.8N/Y17-O/V9 2.9O/Y17-N/V9 2.9 |
| 3f | 2BJU\_A | Plasmepsin 2  | N/D214-O/L301 3.1O/T217-N/D303 2.8O/S218-CD/P304 3.0 (2.2) 136O | O/D214-CA/G302 3.3 (2.3) 155OO/G302-N/F305 2.9 | OD1/D214-N/G216 3.0O/D214-N/T217 3.4 |  |
| 4a | 3QVC\_A, p | HAP zymogen | N/H32-O/F121 3.3O/H32-N/L123 2.7O/S35-N/G124 2.7 | O/H32-CA/G122 3.5 (2.7) 138O | ND1/H32-N/A34 3.4O/H32-N/S35 4.3 | O/V12-N/E86p 3.0N/S14-O/N84p 2.9O/S14-N/N84p 2.9 |
| 4b | 3QVC\_A | HAP zymogen | N/D215-O/L301 3.0O/T218-N/D303 2.9O/S219-CD/P304 3.5 (2.6) 140O | O/D215-CA/G302 3.4 (2.3) 155OO/G302-N/F305 3.0 | OD1/D215-N/A217 3.0O/D215-N/T218 3.5 |  |
| 4e | 3QVI\_A, B | HAP protein  | N/H32-O/F121 3.0O/H32-N/L123 2.7O/S35-N/G124 2.9 | O/H32-CA/G122 3.7 (2.9) 131O | ND1/H32-N/A34 3.0O/H32-N/S35 3.6 | N/L13\_A-O/K7\_B 3.1O/F15-N/K7\_B 2.6 |
| 4f | 3QVI\_A | HAP protein  | N/D215-O/L301 3.0O/T218-N/D303 2.9O/S219-CD/P304 3.7 (3.0) 124O | O/D215-CA/G302 3.4 (2.3) 155OO/G302-N/F305 3.1 | OD1/D215-N/A217 3.0O/D215-N/T218 3.5 |  |
| 5a | 5N7N\_A, p | Procathepsin D  | N/D58-O/L147 3.1O/D58-N/L149 2.8O/S61-N/A150 2.8 | O/D58-CA/G148 3.4 (2.5) 140O | OD1/D58-N/G60 2.9O/D58-N/S61 3.4 | O/V39-N/F9p 3.0N/Y41-O/T7p 3.0O/Y41-N/T7p 2.9 |
| 5b  | 5N7N\_A | Procathepsin D  | N/N249-O/L336 2.9O/T252-N/D338 2.9O/S253-N/V339 3.0 | O/N249-CA/G337 3.5 (2.5) 169OO/G337-N/F340 3.1 | OD1/N249-N/G251 2.8O/N249-N/T252 3.5 |  |
| 4c | 5N71\_A | Cathepsin D  | N/D58-O/L147 3.1O/D58-N/L149 2.9O/S61-N/A150 2.9 | O/D58-CA/G148 3.4 (2.5) 143O | OD1/D58-N/G60 2.8O/D58-N/S61 3.4 | O/V39-N/L35 2.8N/Y41-O/V33 3.1O/Y41-N/V33 2.7 |
| 5d | 5N71\_A | Cathepsin D  | N/N249-O/L336 2.9O/T252-N/D338 2.9O/S253-N/V339 3.1 | O/N249-CA/G337 3.3 (2.3) 164OO/G337-N/F340 2.9 | OD1/N249-N/G251 2.8O/N249-N/T252 3.5 |  |
| 5e | 5N7Q\_A | Cathepsin D  | N/D36-O/L125 2.9O/D36-N/L127 2.8O/S39-N/A128 2.8 | O/D36-CA/G126 3.4 (2.4) 143O | OD1/D36-N/G38 3.0O/D36-N/S39 3.4 | O/V17-N/L13 3.0N/Y19-O/V11 3.1O/Y19-N/V11 2.8 |
| 5f | 5N7Q\_A | Cathepsin D  | N/D227-O/L314 2.9O/T230-N/D316 2.8O/S231-N/V317 3.1 | O/D227-CA/G315 3.3 (2.2) 159OO/G315-N/F318 2.9 | OD1/D227-N/G229 2.9O/D227-N/T230 3.1 |  |
| 6a | 1MIQ\_A, p | Proplasmepsin | N/D34-O/L124 3.1O/D34-N/L126 2.5O/S37-N/G127 2.7 | O/D34-CA/G125 3.4 (2.5) 135O | OD1/D34-N/G36 3.3O/D34-N/S37 3.8 | O/I14-N/E86p 3.1N/F16-O/K84p 2.9O/F16-N/K84p 3.0 |
| 6b | 1MIQ\_A | Proplasmepsin | N/D214-O/L301 3.0O/T217-N/D303 2.8O/T218-CD/P304 3.5 (2.6) 141O | O/D214-CA/G302 3.3 (2.2) 156OO/G302-N/F305 3.2 | OD2/D214-N/G216 2.8O/D214-N/T217 2.9 |  |
| 6e | 1QS8\_A | Plasmepsin  | N/D34-O/L124 3.0O/D34-N/L126 2.9O/S37-N/G127 2.7 | O/D34-CA/G125 3.5 (2.6) 142O | OD2/D34-N/G36 3.0O/D34-N/S37 3.3 | O/M15-N/V11 2.9N/Y17-O/D9 3.2O/Y17-N/D9 2.9 |
| 6f | 1QS8\_A | Plasmepsin  | N/D214-O/L301 3.0O/T217-N/D303 2.7O/T218-CD/P304 3.7 (2.9) 127O | O/D214-CA/G302 3.3 (2.2) 157OO/G302-N/F305 3.2 | OD2/D214-N/G216 2.6O/D214-N/T217 3.3 |  |
| 7a | 5JOD\_A, p | Proplasmepsin 4 | N/D34-O/L124 3.0O/D34-N/L126 2.7O/S37-N/G127 2.8 | O/D34-CA/G125 3.5 (2.6) 140O | OD1/D34-N/G36 2.9O/D34-N/S37 3.7 | O/L14-N/D87p 3.1N/F16-O/K85p 2.7O/F16-N/K85p 3.0 |
| 7b | 5JOD\_A | Proplasmepsin 4 | N/D214-O/L301 3.1O/T217-N/D303 2.8O/S218-CD/P304 3.4 (2.4) 151O | O/D214-CA/G302 3.2 (2.2) 158OO/G302-N/F305 3.0 | OD1/D214-N/G216 2.9O/D214-N/T217 3.2 |  |
| 7e | 1LS5\_A | Plasmepsin 4  | N/D34-O/L124 3.1O/D34-N/L126 2.8O/S37-N/G127 3.2 | O/D34-CA/G125 3.3 (2.5) 128O | OD1/D34-N/G36 2.6O/D34-N/S37 4.3 | O/M15-N/V11 3.0N/Y17-O/D9 3.5O/Y17-N/D9 2.6 |
| 7f | 1LS5\_A | Plasmepsin 4  | N/D214-O/L301 2.8O/T217-N/D303 3.3O/S218-CD/P304 3.7 (2.9) 134O | O/D214-CA/G302 3.2 (2.2) 142OO/G302-N/F305 3.4 | OD1/D214-N/G216 2.5O/D214-N/T217 3.6 |  |
| 8a | 1QDM\_A, p | Prophytepsin | N/D36-O/L126 2.8O/D36-N/L128 2.7O/S39-N/G129 2.8 | O/D36-CA/G127 3.5 (2.6) 139O | OD1/D36-N/G38 3.1O/D36-N/S39 3.4 | O/Q17-N/R13p 2.7N/F19-O/K11p 2.8O/F19-N/K11p 2.8 |
| 8b | 1QDM\_A | Prophytepsin | N/D223-O/L313 2.8O/T226-N/D315 2.8O/S227-N/V316 3.2 | O/D223-CA/G314 2.9 (2.0) 135OO/G314-N/F317 2.7 | OD1/D223-N/G225 2.9O/D223-N/T226 3.9 |  |
| 9a | 1HTR\_B, p | Progastricsin | N/D32-O/M121 3.1O/D32-N/L123 2.9O/S35-N/A124 3.0 | O/D32-CA/G122 3.4 (2.5) 140O | OD1/D32-N/G34 2.9O/D32-N/S35 3.4 | O/A13-N/F10p 2.8N/F15-O/K8p 3.0O/F15-N/K8p 2.9 |
| 9b | 1HTR\_B | Progastricsin | N/D217-O/L304 3.1O/T220-N/D306 2.9O/S221-N/V307 3.0 | O/D217-CA/G305 3.5 (2.5) 165OO/G305-N/F308 2.8 | OD1/D217-N/G219 2.7O/D217-N/T220 3.4 |  |
| 10a | 1TZS\_A, p | Procathepsin E | N/D43-O/L132 3.0O/D43-N/L134 2.7O/S46-N/G135 2.8 | O/D43-CA/G133 3.1 (2.2) 142O | OD1/D43-N/G45 2.9O/D43-N/S46 3.4 | N/F26-O/R9p 2.8O/F26-N/R9p 3.0 |
| 10b | 1TZS\_A | Procathepsin E | N/D228-O/L317 3.1O/T231-N/D319 2.8O/S232-N/V320 3.2 | O/D228-CA/G318 3.4 (2.3) 162OO/G318-N/F321 2.9 | OD1/D228-N/G230 2.7O/D228-N/T231 3.1 |  |
| 11c | 1T6E\_X | Xylanase inhibitor | N/D29-O/A141 3.0O/D29-N/L143 2.8O/G32-N/A144 3.1 | O/D29-CA/G142 3.5 (2.6) 141O | OD1/D29-N/A31 3.2O/D29-N/G32 3.0 | O/L15-N/D10 2.7N/T17-O/T8 2.9O/T17-N/T8 2.9 |
| 11d | 1T6E\_X | Xylanase inhibitor | N/S235-O/L348 2.9O/L238-N/G350 3.0O/P239-N/A351 3.1 | O/S235-CA/G349 3.2 (2.2) 145OO/G349-N/Q352 3.2 | OG/S235-N/R237 3.0O/S235-N/L238 5.1 |  |
| 11e | 1T6G\_A | Xylanase inhibitor | N/D29-O/A141 3.0O/D29-N/L143 2.9O/G32-N/A144 3.0 | O/D29-CA/G142 3.5 (2.6) 143O | OD1/D29-N/A31 3.3O/D29-N/G32 3.0 | O/L15-N/D10 2.8N/T17-O/T8 2.9O/T17-N/T8 2.9 |
| 11f | 1T6G\_A | Xylanase inhibitor | N/S235-O/L348 2.9O/L238-N/G350 2.9O/P239-N/A351 3.1 | O/S235-CA/G349 3.2 (2.2) 145OO/G349-N/Q352 3.1 | OG/S235-N/R237 3.2O/S235-N/L238 5.1 |  |
| 12c | 3AUP\_A | Basic 7S globulin | N/D41-O/A159 3.1O/D41-N/L161 3.0O/G44-N/G162 2.9 | O/D41-CA/G160 3.2 (2.2) 153O | OD1/D41-N/N43 3.3O/D41-N/G44 2.9 | O/L22-N/D17 2.7N/W24-O/Q15 2.9O/W24-N/Q15 2.8 |
| 12d | 3AUP\_A | Basic 7S globulin | N/S265-O/L361 2.9O/T268-N/A363 2.9O/P269-N/R364 2.9 | O/S265-CA/G362 3.3 (2.4) 135OO/G362-N/Q365 2.9 | OG/S265-N/S267 2.9O/S265-N/T268 4.9 |  |
| 13c | 3VLA\_A | EDGP (Fragment) | N/D40-O/A155 3.0O/D40-N/L157 2.9O/G43-N/G158 2.8 | O/D40-CA/G156 3.4 (2.4) 149O | OD1/D40-N/G42 3.4O/D40-N/G43 2.9 | O/Q21-N/D16 2.8N/V23-O/K14 2.9O/V23-N/K14 2.9 |
| 13d | 3VLA\_A | EDGP (Fragment) | N/S271-O/I374 2.9O/N274-N/G376 2.9O/P275-N/H377 2.9 | O/S271-CA/G375 3.6 (2.6) 147OO/G375-N/Q378 3.0 | OG/S271-N/I273 3.0O/S271-N/N274 4.7 |  |
| 13e | 3VLB\_A | EDGP (Fragment) | N/D40-O/A155 3.0O/D40-N/L157 3.2O/G43-N/G158 2.7 | O/D40-CA/G156 3.6 (2.6) 151O | OD1/D40-N/G42 3.7O/D40-N/G43 2.8 | O/Q21-N/D16 2.9N/V23-O/K14 2.8O/V23-N/K14 2.7 |
| 13f | 3VLB\_A | EDGP (Fragment) | N/S271-O/I374 3.1O/N274-N/G376 2.7O/P275-N/H377 2.9 | O/S271-CA/G375 4.0 (3.2) 129OO/G375-N/Q378 3.4 | OG/S271-N/I273 3.5O/S271-N/N274 4.9 |  |
| Family: Retroviral protease (retropepsin) |
| 14c | 3IXO\_A | HIV-1 protease | N/D25-O/I85 3.0O/A28-N/R87 2.7O/D29-N/N88 3.1 | O/D25-CA/G86 3.8 (2.7) 170OO/G86-N/M89 3.0 | OD1/D25-N/G27 2.9O/D25-N/A28 3.1 |  |
| 14d | 3IXO\_ B | HIV-1 protease | N/D25-O/I85 3.0O/A28-N/R87 2.9O/D29-N/N88 3.4 | O/D25-CA/G86 3.9 (2.9) 162OO/G86-N/M89 3.0 | OD1/D25-N/G27 2.8O/D25-N/A28 3.2 |  |
| 14e | 5YOK\_A | HIV-1 protease | N/D25-O/I85 2.9O/A28-N/R87 2.8O/D29-N/N88 3.3 | O/D25-CA/G86 3.7 (2.7) 167OO/G86-N/L89 3.1 | OD1/D25-N/G27 2.9O/D25-N/A28 3.2 |  |
| 14f | 5YOK\_ B | HIV-1 protease | N/D25-O/I85 2.9O/A28-N/R87 2.8O/D29-N/N88 3.3 | O/D25-CA/G86 3.7 (2.7) 170OO/G86-N/L89 3.1 | OD1/D25-N/G27 2.7O/D25-N/A28 3.1 |  |
| 15c | 3NR6\_A | XMRV protease | N/D32-O/L93 2.9O/A35-N/R95 2.8O/Q36-N/D96 3.1 | O/D32-CA/G94 3.5 (2.5) 163OO/G94-N/L97 3.2 | OD1/D32-N/G34 2.9O/D32-N/A35 3.0 |  |
| 15d | 3NR6\_B | XMRV protease | N/D32-O/L93 2.9O/A35-N/R95 2.8O/Q36-N/D96 3.1 | O/D32-CA/G94 3.5 (2.5) 166OO/G94-N/L97 3.2 | OD1/D32-N/G34 3.0O/D32-N/A35 3.1 |  |
| 15e | 3SLZ\_A | XMRV protease | N/D32-O/L93 2.8O/A35-N/R95 2.9O/Q36-N/D96 3.7 | O/D32-CA/G94 3.4 (2.3) 168OO/G94-N/L97 3.1 | OD1/D32-N/G34 3.0O/D32-N/A35 3.1 |  |
| 15f | 3SLZ\_B | XMRV protease | N/D32-O/L93 2.8O/A35-N/R95 2.9O/Q36-N/D96 3.8 | O/D32-CA/G94 3.4 (2.4) 166OO/G94-N/L97 3.1 | OD1/D32-N/G34 2.9O/D32-N/A35 3.1 |  |
| Family: Dimeric aspartyl proteases |
| 16c | 4Z2Z\_A | DDI1 protease  | N/D220-O/I289 2.9O/A223-N/L291 2.8O/Q224-N/D292 3.1 | O/D220-CA/G290 3.6 (2.6) 166OO/G290-N/M293 3.3 | OD1/D220-N/G222 2.9O/D220-N/A223 3.2 |  |
| 16d | 4Z2Z\_B | DDI1 protease  | N/D220-O/I289 2.9O/A223-N/L291 2.9O/Q224-CB/D292 3.9 (3.1) 136O | O/D220-CA/G290 3.5 (2.5) 154OO/G290-N/M293 3.3 | OD1/D220-N/G222 2.9O/D220-N/A223 3.2 |  |
| 17c | 5C9F\_A | ApRick protease | N/D140-O/L209 3.0O/A143-N/M211 2.9O/S144-N/S212 2.9 | O/D140-CA/G2103.4 (2.3) 162OO/G210-N/L213 3.4 | OD1/D140-N/G142 3.1O/D140-N/A143 3.1 |  |
| Family: LPG0085-like |
| 18c | 2PMA\_A | Protein Lpg0085 | N/D47-O/L145 2.9O/A50-N/R147 2.8O/K51-N/D148 3.1 | O/D47-CA/G146 3.5 (2.5) 154OO/G146-N/A149 3.2 | OD1/D47-N/G49 3.0O/D47-N/A50 3.4 |  |
| 18d | 2PMA\_I | Protein Lpg0085 | N/D47-O/L145 2.8O/A50-N/R147 2.8O/K51-N/D148 3.1 | O/D146-CA/G146 3.4 (2.4) 158OO/G146-N/A149 3.2 | OD1/D47-N/G49 2.9O/D47-N/A50 3.4 |  |

**Table S3**. Conserved geometric parameters (distance and angle) of contacts between hydrolase and ligand in 9 acid proteases pepsin-like and retroviral protease (retropepsin) families.

|  |  |  |  |
| --- | --- | --- | --- |
| N | PDB ID | Protein  | D-loops / ligand  |
| Superfamily: Acid proteases |
| Family: Pepsin-like |
| 1e | 6XCZ\_A | Pepsin  | OD1/D32-O2/ROC401 2.4 | O/G34-N2/ROC401 3.0  | N/S36-HOH527 2.8HOH527-OD1/ROC401 2.6 |
| 1f | 6XCZ\_A | Pepsin  | OD2/D215-O2/ROC401 2.8 | O/G217-C21/ROC401 3.3 (2.4) 141OO/G217-CM/ROC401 3.3 (2.5) 131O | N/S219-HOH645 3.3HOH645-N3/ROC401 3.2 |
| 2e | 3K1W\_A | Renin  | OD1/D38-N27/BFX342 3.0 | O/G40-C24/BFX342 3.5 (2.5) 148O | OG/S41-C17/BFX3423.9 (2.9) 160O |
| 2f | 3K1W\_A | Renin  | OD2/D226-N27/BFX342 2.7 | O/G228-C34/BFX342 3.2 (2.3) 135O | OG/S230-C37/BFX3423.5 (2.9) 117O |
| 3e | 2BJU\_A | Plasmepsin 2  | OD1/D34-C5/IH41330 3.9 (2.8) 159O | O/G36-C6/IH41330 3.9 (3.1) 134O | N/A |
| 3f | 2BJU\_A | Plasmepsin 2  | OD2/D214-N13/IH41330 2.6 | O/G216-C32/IH41330 3.2 (2.8) 106OO/G216-C22/IH41330 3.9 (2.9) 161O | N/S218-O1/IH41330 2.9 |
| 4e | 3QVI\_A | HAP protein  | NE2/H32-HOH338 3.1HOH338-CBB/K95329 3.4 (2.6) 129O | O/A34-OAI/K95329 2.9 | OG/S35-OAG/K95329 3.0 |
| 4f | 3QVI\_A, B | HAP protein  | OD2/D215\_A-HOH364 2.5HOH364-OAI/K95329 2.9 | OE2/E278A\_B-NBD/K95329 2.7 | OE2/E278A\_B-NBD/K95329 2.7 |
| 5e | 5N7Q\_A | Cathepsin D  | OD1/D36-OH/STA504 2.6 | O/G38-N/ALA505 2.9 | N/A40-HOH735 2.8HOH735-CB/ALA505 3.9 (3.5) 104O |
| 5f | 5N7Q\_A | Cathepsin D  | OD2/D227-OH/STA504 2.6 | O/G229-N/STA504 3.1 | N/S231-O/VAL502 2.9 |
| 6e | 1QS8\_A | Plasmepsin  | OD2/D34-OH/STA404 2.5 | O/G36-N/ALA405 3.1 | N/A38-HOH349 2.8HOH349-CB/ALA405 3.9 (3.5) 103O |
| 6f | 1QS8\_A | Plasmepsin  | OD1/D214-OH/STA404 2.4 | O/G216-N/STA404 3.5 | N/T218-O/VAL402 3.0 |
| 7e | 1LS5\_A | Plasmepsin 4  | OD1/D34-OH/STA4 2.3 | O/G36-N/ALA5 2.9 | CB/S37-CB/ALA5 3.6 |
| 7f | 1LS5\_A | Plasmepsin 4  | OD1/D214-OH/STA4 3.1 | O/G216-N/STA4 2.7 | N/S218-O/VAL2 2.8 |
| Family: Retroviral protease (retropepsin) |
| 14f | 5YOK\_ A | HIV-1 protease | OD1/D25-O2/8Z0101 2.9 | O/G27-HOH308 3.0HOH308-N5/8Z0101 2.9 | N/D29-HOH308 2.9HOH308-N5/8Z0101 2.9 |
| 14f | 5YOK\_B | HIV-1 protease | OD1/D25-O2/8Z0101 2.6 | O/G27-N3/8Z0101 3.3 | N/D29-O3/8Z0101 3.0 |
| 15e | 3SLZ\_A | XMRV protease | OD1/D32-O51/3TL126 2.5 | O/G34-N1/3TL126 2.9 | N/Q36-O4/3TL126 2.9OE1/Q36-N4/3TL126 3.1 |
| 15f | 3SLZ\_B | XMRV protease | OD1/D32-O51/3TL126 2.8 | O/G34-N51/3TL126 3.0 | N/Q36-O54/3TL126 3.0OE1/Q36-N54/3TL126 3.1 |