Supporting Information for

Covalent Labeling with Diethylpyrocarbonate is Sensitive to Residue Microenvironment, Providing Improved Analysis of Protein Higher Order Structure by Mass Spectrometry

Patanachai Limpikirati, Xiao Pan, and Richard W. Vachet*

Department of Chemistry, University of Massachusetts Amherst, Amherst, Massachusetts 01003, United States

* Corresponding author

AUTHOR INFORMATION

Patanachai Limpikirati: plimpikirati@umass.edu

Xiao Pan: xpan@umass.edu

Richard W. Vachet: rwvachet@chem.umass.edu

CONTENT

SUPPLEMENTAL EXPERIMENTAL SECTION

A. Materials

Human full-length beta-2-microglobulin (β2m) (#126-11) and recombinant human growth hormone (hGH) (#4769) were obtained from Lee Biosolutions (Maryland Heights, MO) and BioVision (Milpitas, CA), respectively. Bovine ubiquitin (#U6253), bradykinin (#B3259), diethylpyrocarbonate (DEPC) (#D5758), imidazole (#I5513), iodoacetamide (#I6125), tris(2-carboxyethyl)phosphine (TCEP) (#C4706), 3-morpholinopropane-1-sulfonic acid (MOPS) (#M1254), and MOPS sodium salt (#M9381) were all purchased from Sigma-Aldrich (St. Louis, MO). Preproenkephalin was ordered from American Peptide Company (Sunnyvale, CA). Model peptides (Fmoc-DGXGG-amide, where X = H, K, Y, S, or T) were custom synthesized by GenScript USA Inc (Piscataway, NJ). Pierce™ sulfo-NHS-Acetate (#26777) was obtained from Thermo Scientific (Waltham, MA). Sodium phosphate monobasic monohydrate (#S0710) was purchased from EM Science (Darmstadt, Hesse, Germany). Sodium phosphate dibasic anhydrous (#S374), N,N-Dimethylformamide (#D119), formic acid (#A117), acetonitrile (#A998), and water (#W7) were obtained from Fisher Scientific (Fair Lawn, NJ). Immobilized trypsin (#EN-251) and chymotrypsin (#EN-261) were purchased from Princeton Separations (Adelphia, NJ). Amicon® centrifugal filters (#UFC5010 and #UFC5003) were bought from EMD Millipore (Burlington, MA). All reagents used in this study have no known potential hazards.

B. DEPC Labeling and LC-MS/MS Analyses of Model Peptides

DEPC Labeling Reactions. A lyophilized powder of model peptide (Fmoc-DGXGG-amide, where X = H, K, Y, S, or T) was first reconstituted in water (His and Lys) or dimethylformamide (Ser, Thr, and Tyr), and the resulting solution of peptide was then diluted in 10 mM MOPS buffer (pH 7.4) to make a final solution. Each model peptide (50 μM) was reacted with DEPC at 37 °C for 5 min at a DEPC to protein molar ratio of 5 to 1 (His, Lys, and Tyr) or 50 to 1 (Ser and Thr). The reaction was then quenched by the addition of imidazole at a 1:50 DEPC to imidazole molar ratio.

Bradykinin and preproenkephalin peptides were reconstituted in water. A peptide mixture was prepared in 10 mM MOPS buffer (pH 7.4) containing 10 μM of each peptide. The N-termini

of the peptides were acetylated before labeling them with DEPC. The N-terminal blocking was initiated by adding sulfo-NHS-acetate in a molar excess of 600, and the solution was reacted for 1 h at room temperature. Subsequently, the peptide mixture was reacted with DEPC at 37 °C for 5 min at a molar excess of 50. The reaction was then quenched by the addition of imidazole.

Liquid Chromatography (LC). Following quenching online LC-MS/MS analyses were conducted on a Thermo Scientific Ultimate 3000 HPLC (Waltham, MA). LC separation was performed on a Thermo Scientific Acclaim™ PepMap™ RSLC C18 column (15 cm x 300 μm, 2 μm particle size). LC/MS-grade water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as mobile phases. A flow rate of 4 μL/min was used, and 1 μL of sample was first loaded and desalted at 5%B during the first 5 min. An isocratic elution mode at 40%B over 15 min was then applied to separate unmodified and modified peptides.

For a mixture of bradykinin and preproenkephalin peptides, 5 μL of sample was first loaded and desalted during the first 5 min. After that a linear gradient was increased from 5%B to 50%B over 50 min to separate unmodified and modified peptides.

Mass Spectrometry (MS). Mass spectra from the online LC-MS/MS were acquired on a Bruker AmaZon quadrupole ion trap (Billerica, MA). The electrospray needle voltage was operated using a positive mode at ~4 kV, and the capillary temperature was set to 300 ^{O}C. Tandem spectra were collected for the top intense species with ion abundances above 1,000 from each mass spectrum. Collisional-induced dissociation (CID) was conducted with a ramp energy 60% to 180% of 1.5 V amplitude.

Peptide Identification and Peak Quantification. Bruker Compass™ Data Analysis software was used to reconstruct extracted ion chromatograms (XICs) of unmodified and modified peptides. Peptide identification and peak quantification were performed in a manual manner using tandem spectra and mass spectral peak areas, respectively. DEPC modification levels (%CL) of each labeled residue were calculated using **Eq. S1** (see **Section D. Determination of Modification Percentages**).

C. LC-MS/MS Analyses of Protein Digests (Using a Thermo Scientific Dionex Ultimate 3000 HPLC and a Bruker AmaZon Quadrupole Ion Trap Mass Spectrometer)

Liquid Chromatography (LC). Chymotryptic digests obtained from the DEPC reactions on β2m intact proteins were analyzed using a Thermo Scientific Dionex Ultimate 3000 HPLC (Waltham, MA). 5 μL of sample was loaded and separation of peptides was performed on a Thermo Scientific Acclaim™ PepMap™ RSLC C18 column (15 cm x 300 μm, 2 μm particle size). A flow rate of 4 μL/min was used, and desalting was performed at 5%B during the first 5 min after sample injection. A 50-min linear gradient was applied with %B increased from 5%B to 50%B to separate peptides. The gradient was finally elevated to and held at 95%B to flush a column.

Mass Spectrometry (MS). Mass spectra were acquired on a Bruker AmaZon quadrupole ion trap (Billerica, MA). The electrospray needle voltage was kept at ~4 kV (positive mode), and the capillary temperature was set to 300^{O}C. The top intense peptides with ion abundances above 1,000 from each mass spectrum were selected for MS/MS acquisition. CID was conducted at 1.5 V amplitude with a ramp energy 60% to 180%. Because of the large number of measured peaks, active exclusion of 0.5 min was activated after 2 spectra were acquired for any given precursor ion.

Peptide Identification and Peak Quantification. See Section B.

D. Determination of Modification Percentages

DEPC modification levels (%) of each labeled residue were calculated as follows **(Eq. S1)**.

$$\%\ DEPC\ labeling = \frac{\sum_{i=1}^{n}\sum_{z=1}^{m} A_{i,z}^{modi}}{\sum_{i=1}^{n}\sum_{z=1}^{m} A_{i,z}^{modi} + \sum_{i=1}^{n}\sum_{z=1}^{m} A_{i,z}^{unmodi}} \times 100 \qquad \text{(S1)}$$

where $A_{i,z}^{unmodi}$ is the peak area of DEPC-unmodified peptide whose sequence (i) contains the residue of interest and possesses a certain charge state (z), and $A_{i,z}^{modi}$ is the peak area of peptide in which the residue of interest is DEPC-modified. An illustration of this calculation can be found in **Figure S1**.

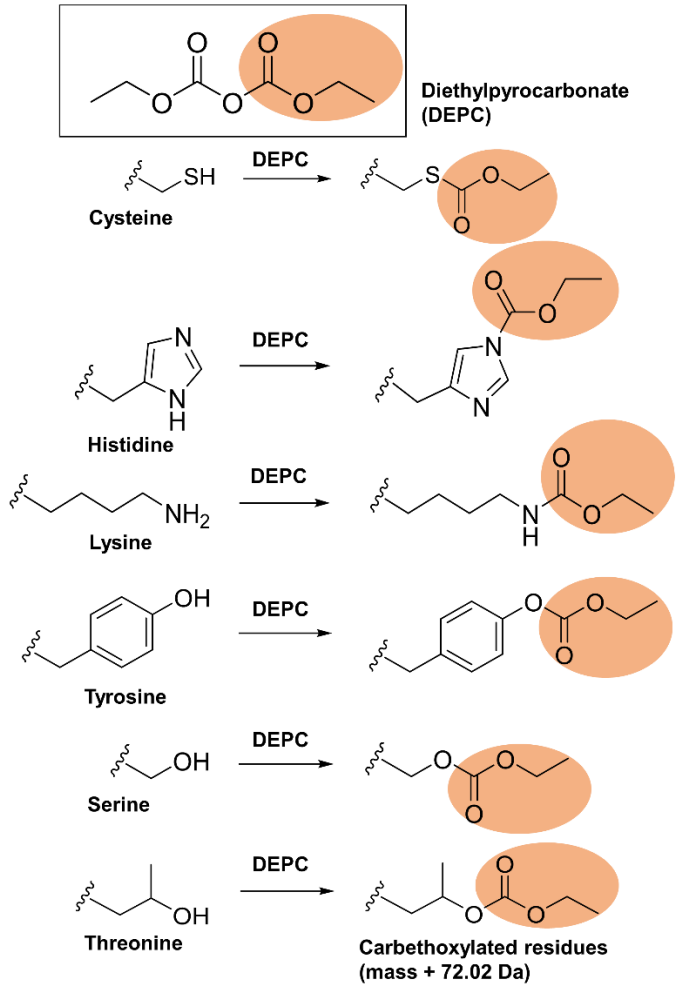
Note that DEPC modification percentages calculated by **Eq. S1** are only used for the relative quantitation, i.e. comparing the DEPC labeling at the same residue under different conditions (labeling on intact protein vs. protein digest). The modification levels do not reflect the absolute quantitation of modified species as the addition of a carbethoxyl group to the modified peptide and different LC solvent conditions during gradient elution of peptides result in different ionization efficiencies of the unmodified and modified peptides.

E. Solvent Accessible Surface Area (SASA) calculation

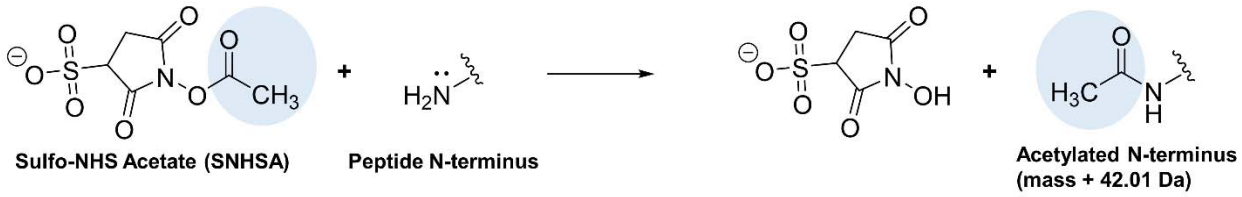
SASA of amino acid side chain was calculated from the Protein Data Bank (PDB) atomic coordinates using GETAREA 1.0 beta.[4] A probe radius of 1.4 Å which represents the van der Waals sphere of water was used in a calculation. Atomic coordinates of the 3D structures of β2m (PDB ID: 1JNJ),[1] ubiquitin (PDB ID: 1UBQ),[2] and hGH (PDB ID: 1HGU)[3] were submitted to calculate individual side-chain SASA. Note that human and bovine ubiquitin have the same amino acid sequence. Even though bovine ubiquitin was used in experiments, the SASA and structural features considered in this study were obtained from the PDB structure of human ubiquitin (PDB ID: 1UBQ). The calculated SASA was compared to the surface area of the side chain in a Gly-X-Gly tripeptide, where X is a side chain of interest, to generate a percent ratio (%SASA).

SUPPLEMENTAL SCHEMES, FIGURES, AND TABLES

Scheme S1. DEPC labeling reactions of Cys, His, Lys, Tyr, Ser, and Thr



Scheme S2. Acetylation reaction of peptide N-terminus



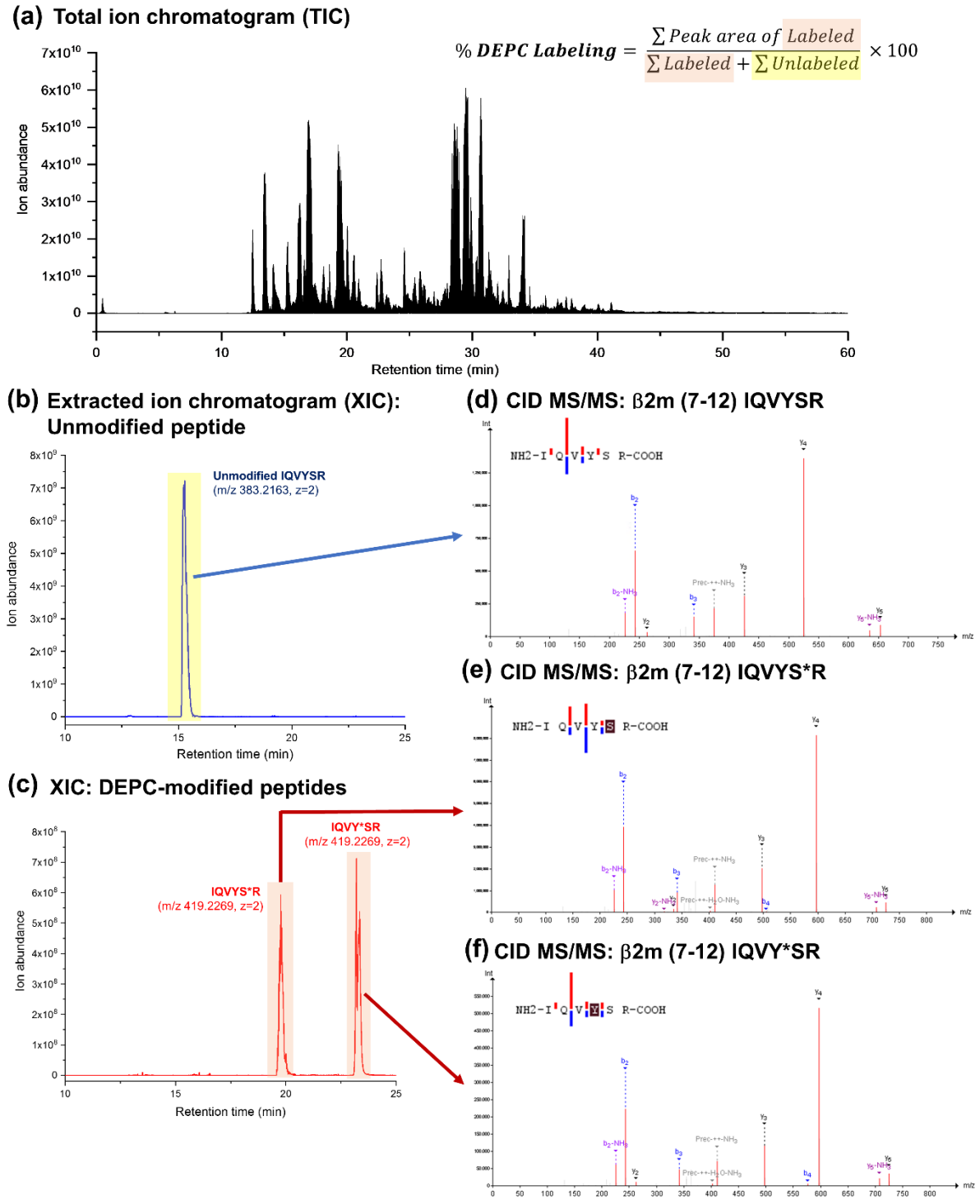


Figure S1. Illustration of how the DEPC modification levels are calculated. After DEPC labeling and proteolytic digestion, (a) LC-MS analysis of the digested protein is performed. Peak areas of (b) unlabeled and (c) labeled peptides in a chromatogram are used to calculate the labeling percentage. During LC-MS, peptides are subjected to CID MS/MS. Tandem mass spectra of (d) unlabeled and (e) & (f) labeled peptides obtained at specific retention time are used for peptide sequencing and/or identification of DEPC labeled site.

DEPC Labeling Reactivity of Weakly Nucleophilic Residues

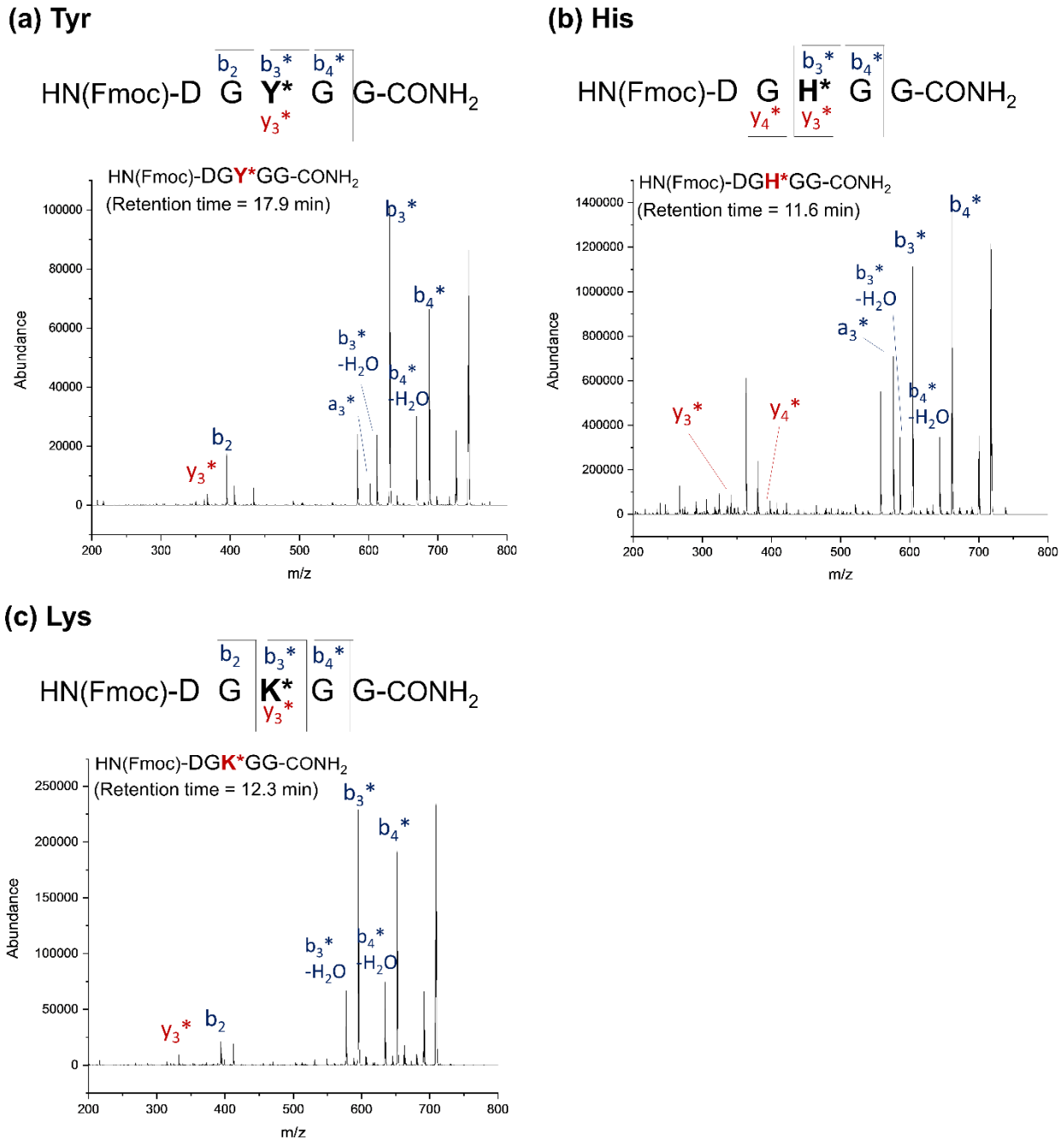


Figure S2. DEPC labeling of model peptides. MS/MS assignments for nucleophilic residues that are modified in model peptides Fmoc-DGXGG-amide, where X is a DEPC modifiable residue (a) Tyr, (b) His, or (c) Lys, after allowing the peptide to react with DEPC at a molar ratio of 5:1 (DEPC:peptide).

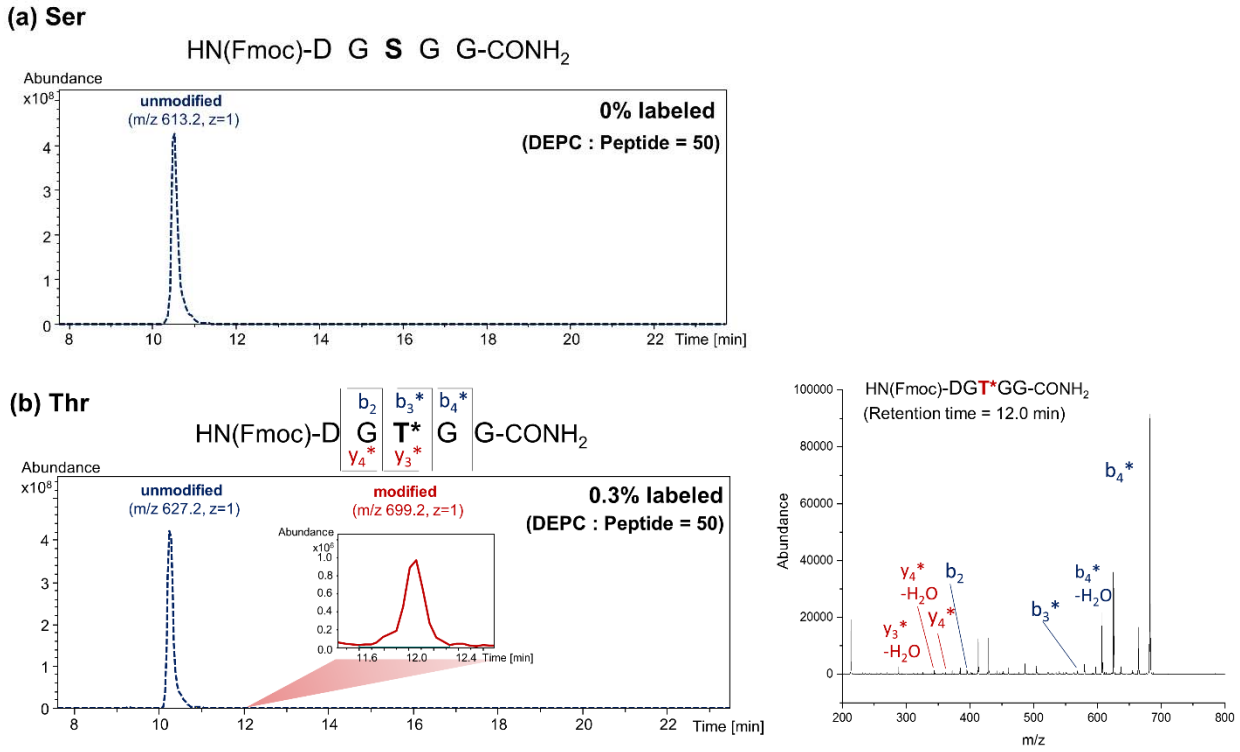


Figure S3. DEPC labeling of model peptides. Extracted ion chromatograms (XICs, left) of the +1 ions of the unmodified and DEPC-modified peptides Fmoc-DGXGG-amide, where X is a DEPC modifiable (a) Ser or (b) Thr, after allowing the peptide to react with DEPC at a DEPC:peptide molar ratio of 50:1. Modification percentages are calculated from peak areas in XIC. Tandem mass spectra (right) are used to confirm the site of modification. MS/MS assignment of the modified peptide is shown above the XICs.

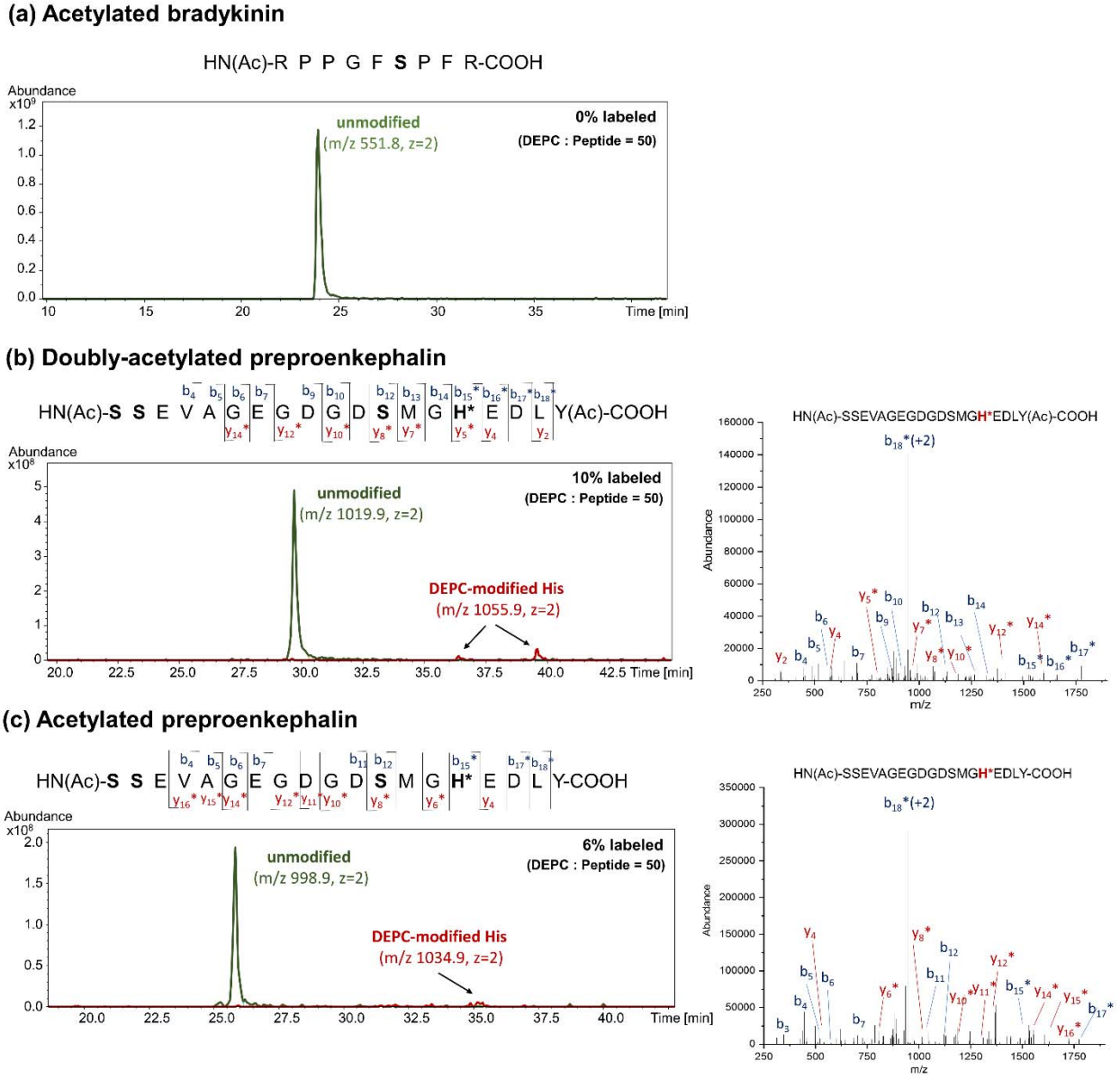


Figure S4. DEPC labeling of bradykinin and preproenkephalin peptides. Extracted ion chromatograms (XICs, left) of the +2 ions of the unmodified and DEPC-modified versions of the N-terminally blocked peptides (a) bradykinin, (b) & (c) preproenkephalin, after reacting the peptide mixture with DEPC at a DEPC to peptide molar ratio of 50 to 1. Tyr and Ser residues in these peptides are unreactive with DEPC even when the other reactive site (N-terminus) are blocked via acetylation with sulfo-NHS-acetate in these peptides **(Scheme S2)**. Tandem mass spectra (right) were used to confirm the site of modification. MS/MS assignments of the modified peptides are shown above the XICs.

Influence of Higher-Order Structure on the Covalent Labeling Based Structural Analysis of Proteins and Identifying Structural Features that Tune the Reactivity of Weak Nucleophiles

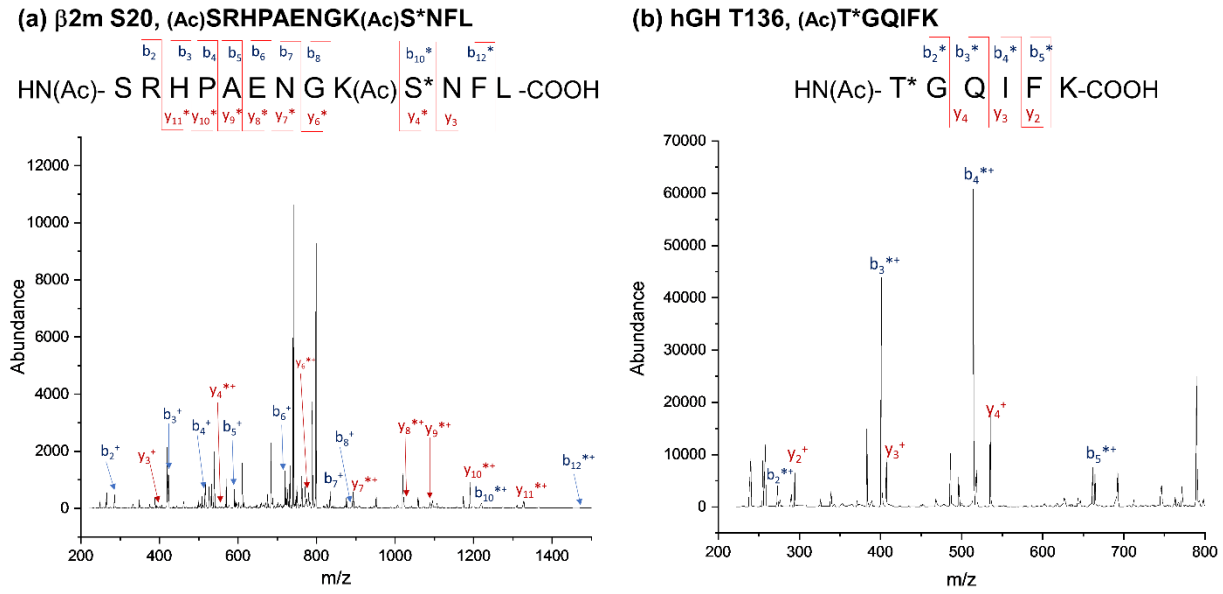


Figure S5. MS/MS assignments for weakly nucleophilic residues that are modified at peptide level. When reacting the N-terminally blocked peptides at DEPC to peptide molar ratios of 4 to 1 or 5 to 1, the reactivity of the Ser, Thr, and Tyr residues in free peptides are found to be lower than in the intact proteins. Only two residues in three model proteins are found to be modified. This Figure S-hows tandem mass spectra acquired after CID of (a) HN(Ac)-SRHPAENGK(Ac)S*NFL-COOH peptide from chymotryptic digest of β-2-microglobulin (β2m), where Ser20 is the DEPC modification site, and (b) HN(Ac)-T*GQIFK-COOH peptide from tryptic digest of human growth hormone (hGH), where Thr136 is the modification site. Product ions with DEPC-modified residue have mass addition of 72.02 Da while acetylated N-terminus or Lys contributes to mass addition of 42.01 Da.

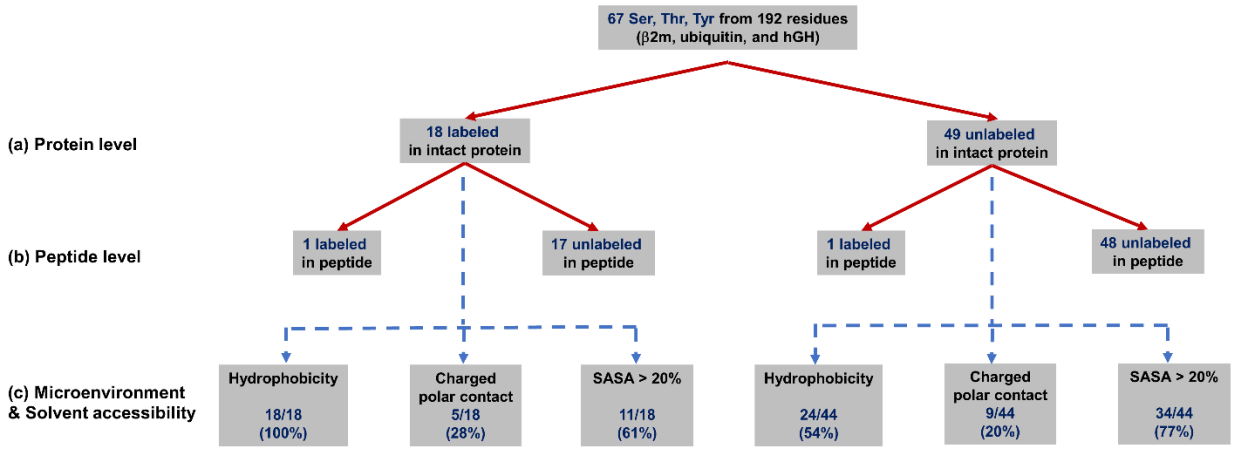


Figure S6. Flow chart summarizing the covalent labeling results and the structural features of weakly nucleophilic residues in β2m, ubiquitin, and hGH. The numbers of these residues that are found to be labeled or unlabeled in (a) intact proteins and (b) peptide fragments are shown in the chart. Structural features that could affect the labeling reactivity of these residues, such as the presence of nearby hydrophobic residues or charged residues and the solvent accessible surface area (SASA) above 20%, are indicated in (c). The fraction of labeled or unlabeled residues that have each specific structural feature is shown in the chart. For the unlabeled Ser, Thr, and Tyr residues, structural features are investigated for 44 out of 49 residues as 5 of the residues in hGH are not resolved in the crystal structure (PDB ID: 1HGU).[3]

Scheme S3. Molecular schemes showing how both positively-charged and negatively-charged residues could affect Ser, Thr, and Tyr. (a) Deprotonated forms of these weak nucleophiles can presumably be stabilized by nearby positively-charged residues (Arg, Lys), increasing their nucleophilicity, and (b) the hydroxyl group of Ser, Thr, and Tyr side chains can form an ionic hydrogen bond with nearby negatively-charged residues (Asp, Glu), which might also increase their nucleophilicity.

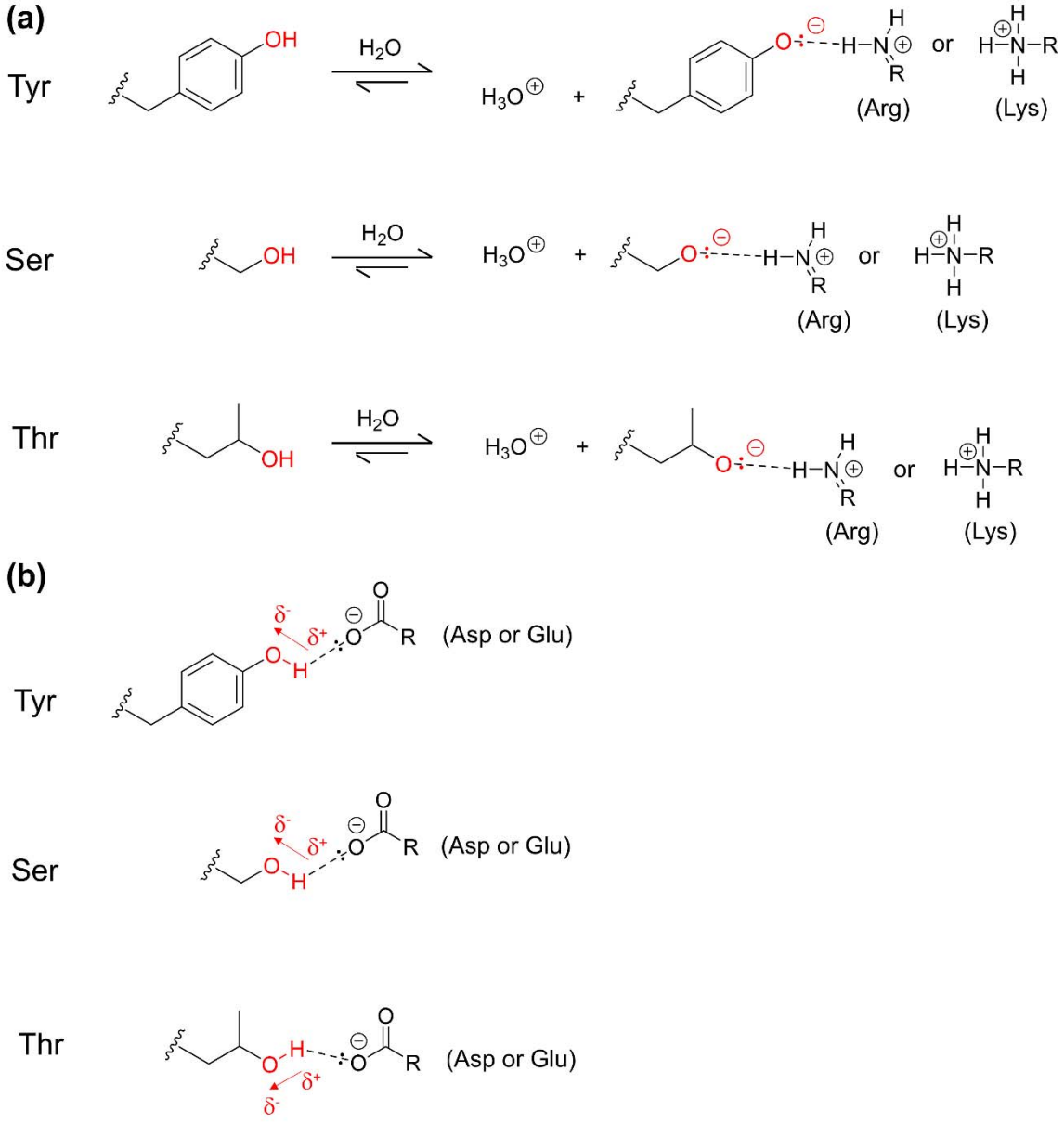


Table S1. DEPC modification percentages of nucleophilic residues in β2m intact protein and its proteolytic peptides. Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1JNJ. ✓ and ✕ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

β2m residue	Secondary structure	pK_a ‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
T4	Random coil	N/A[a]	81.5	✓	✕ **- Overall: Uncharged, nothing**	✓ - OH close to P5, L87	59 ± 5 (T4 & N-term)	N.D.[c]	0.02 ± 0.01
K6	β-sheet	10.42	79.6	✓	✕ - H-bonding with backbone CO of D8 **- Overall: Uncharged, polar contact**	✕	41 ± 5	N/A[b]	N/A
Y10	β-sheet	10.98	41.6	✓	✕ - H-bonding with side chain $CONH_2$ of N24 **- Overall: Uncharged, polar contact**	✓ - OH close to Y26, L65	15 ± 2	N.D.	3.9 ± 0.8
S11	β-sheet	N/A	10.3	✕	✕ - H-bonding with side chain aromatic amine of W95 - H-bonding with backbone CO of F22, R97, M99 - H-bonding with backbone NH of R12, D98, M99 **- Overall: Uncharged, polar contact**	✓ - OH close to A15, L23, F70, M99	3.0 ± 0.5	N.D.	0.02 ± 0.01
H13	Random coil	6.09	47.7	✓	✕ **- Overall: Uncharged, nothing**	✓ - imidazole amine close to F22	30 ± 3	N/A	N/A
K19	Random coil	10.42	77.7	✓	✓ - Salt bridge with side chain COOH of E16 - H-bonding with backbone CO of S20, N21 **- Overall: Charged, (-)**	✕	30 ± 2	N/A	N/A
S20	Random coil	N/A	61.6	✓	✓ - H-bonding with side chain OH of T71 - H-bonding with backbone CO of E69 - H-bonding with side chain COOH of E69 **- Overall: Charged, (-)**	✕	N.D.	2.6 ± 0.6	N.D.

[a] N/A = not applicable; [b] N-terminal blocking using SNHSA also results in side reaction (acetylation) at Lys and His residues in free peptides, so they are not discussed here.
[c] N.D. = not detected (less than DEPC labeling threshold of 0.01%[5])
* SASA calculated using GETAREA[4], as explained in detail in the Supplemental Experimental Section; ‡ pKa calculated using PROPKA[6]

β2m residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
Y26	β-sheet	10.20	26.4	✓	✗ **- Overall: Uncharged, nothing**	✓ - OH close to L65	N.D.	N.D.	N.D.
S28	β-sheet	N/A	18.2	✗	✗ - H-bonding with side chain $CONH_2$ of Q8 - H-bonding with side chain OH of Y26 - H-bonding with backbone NH of G29 **- Overall: Uncharged, polar contact**	✓ - OH close to Y26	N.D.	N.D.	N.D.
H31	Random coil	5.98	33.5	✓	✗ **- Overall: Uncharged, nothing**	✓ - imidazole amine close to I1	5 ± 3	N/A	N/A
S33	Random coil	N/A	70.1	✓	✗ **- Overall: Uncharged, nothing**	✓ - OH close to P32, F62	1.6 ± 0.4	N.D.	N.D.
K41	Random coil	12.73	12.4	✗	✓ - Salt bridge with side chain COOH of E44, D76 - H-bonding with side chain OH of Y78 - H-bonding with backbone CO of E44 **- Overall: Charged, (-)**	✓ - NH_2 close to I46	1.7 ± 0.8	N/A	N/A
K48	Random coil	10.62	87.0	✓	✓ - Salt bridge with side chain COOH of E47 **- Overall: Charged, (-)**	✗	5 ± 2	N/A	N/A
H51	Random coil	7.15	57.5	✓	✓ - Salt bridge with side chain COOH of D53 - H-bonding with backbone NH of D53 **- Overall: Charged, (-)**	✓ - imidazole amine close to L64	4 ± 1	N/A	N/A
S52	Random coil	N/A	50.7	✓	✗ - H-bonding with side chain OH of Y67 - H-bonding with backbone CO of D53, L65 - H-bonding with backbone NH of D53, L65 **- Overall: Uncharged, polar contact**	✓ - OH close to L65, Y67	N.D.	N.D.	N.D.
S55	Random coil	N/A	55.3	✓	✗ - H-bonding with backbone CO of S57 **- Overall: Uncharged, polar contact**	✓ - OH close to L54	0.6 ± 0.3	N.D.	N.D.

β2m residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
S57	Random coil	N/A	34.9	✓	✓ - H-bonding with side chain COOH of D59 - H-bonding with side chain OH of S61 - H-bonding with backbone CO of F56, S61 - H-bonding with backbone NH of K58, D59 **- Overall: Charged, (-)**	✓ - OH close to F56	1.8 ± 0.9	N.D.	N.D.
K58	Random coil	10.29	88.5	✓	✗ **- Overall: Uncharged, nothing**	✓ - NH_2 close to F56	9 ± 2	N/A	N/A
S61	Random coil	N/A	28.1	✓	✓ - H-bonding with side chain guanidinium of R3 - H-bonding with backbone NH of G29, S57, F62 - H-bonding with backbone CO of G29, F56, S57, D58 **- Overall: Charged, (+)**	✓ - OH close to F56, W60, F62, Y63	10 ± 1	N.D.	N.D.
Y63	Random coil	9.54	27.7	✓	✗ - H-bonding with side chain OH of Y26, S55 **- Overall: Uncharged, polar contact**	✓ - OH close to Y26, F56	0.3 ± 0.2	N.D.	N.D.
Y66	β-sheet	13.25	7.5	✗	✗ **- Overall: Uncharged, nothing**	✓ - OH close to V27, F30, I35, V37, L64	0.3 ± 0.2	N.D.	N.D.
Y67	β-sheet	10.46	22.8	✓	✓ - H-bonding with side chain OH of S52 - H-bonding with backbone CO of H51 - H-bonding with side chain COOH of E50 **- Overall: Charged, (-)**	✓ - OH close to L65	N.D.	N.D.	1.3 ± 0.6
T68	β-sheet	N/A	0.1	✗	✓ - H-bonding with backbone CO of K48 - H-bonding with backbone NH of E69 - H-bonding with side chain COOH of E69 **- Overall: Charged, (-)**	✓ - OH close to L39, I46, V49, F70	N.D.	N.D.	N.D.

β2m residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
T71	Random coil	N/A	29.8	✓	✗ - H-bonding with side chain OH of S20 - H-bonding with backbone CO of E69 **- Overall: Uncharged, polar contact**	✗	N.D.	N.D.	N.D.
T73	Random coil	N/A	63.9	✓	✓ - H-bonding with side chain COOH of D76 - H-bonding with backbone NH of E74, K75, D76 **- Overall: Charged, (-)**	✗	N.D.	N.D.	N.D.
K75	Random coil	10.47	100	✓	✓ - Salt bridge with side chain COOH of E74, E77 **- Overall: Charged, (-)**	✗	2.3 ± 0.9	N/A	N/A
Y78	Random coil	10.34	5.1	✗	✓ - H-bonding with side chain NH_2 of K41 - H-bonding with backbone NH of T71 - H-bonding with backbone CO of T71 **- Overall: Charged, (+)**	✓ - OH close to I46, F70	N.D.	N.D.	N.D.
H84	Random coil	5.78	0.0	✗	✗ - H-bonding with backbone NH of T86 **- Overall: Uncharged, polar contact**	✓ - imidazole amine close to F30, V85	2.0 ± 0.3	N/A	N/A
T86	Random coil	NA	9.2	✗	✗ - H-bonding with backbone NH of L87 - H-bonding with backbone CO of V85 **- Overall: Uncharged, polar contact**	✓ - OH close to V85, L87	N.D.	N.D.	N.D.
S88	Random coil	NA	96.7	✓	✗ - H-bonding with backbone CO of V85 **- Overall: Uncharged, polar contact**	✓ - OH close to L87	N.D.	N.D.	N.D.
K91	Random coil	10.36	71.3	✓	✗ - H-bonding with side chain $CONH_2$ of Q89 - H-bonding with backbone CO of I92 **- Overall: Uncharged, polar contact**	✓ - NH_2 close to V93	4 ± 3	N/A	N/A
K94	Random coil	10.56	80.3	✓	✓ - Salt bridge with side chain COOH of E77 - H-bonding with backbone CO of W95 **- Overall: Charged, (-)**	✓ - NH_2 close to I92, M99	6 ± 2	N/A	N/A

Table S2. DEPC modification percentages of nucleophilic residues in ubiquitin (Ub) intact protein and its proteolytic peptides. Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1UBQ. ✓ and ✕ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

Ub residue	Secondary structure	pK_a ‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
K6	β-sheet	10.37	59.7	✓	✕ **- Overall: Uncharged, nothing**	✕	0.8 ± 0.6	N/A[b]	N/A
T7	β-sheet	N/A[a]	14.2	✕	✕ - H-bonding with backbone CO of T9, K11 - H-bonding with backbone NH of T9, K11 **- Overall: Uncharged, polar contact**	✓ - OH close to L8	13 ± 2	N.D.[c]	N.D.
T9	Random coil	N/A	85.5	✓	✕ - H-bonding with side chain OH of T7 **- Overall: Uncharged, polar contact**	✕	N.D.	N.D.	N.D.
K11	β-sheet	11.02	59.8	✓	✓ - Salt bridge with side chain COOH of E34 **- Overall: Charged, (-)**	✓ - NH_2 close to I13	N.D.	N/A	N/A
T12	β-sheet	N/A	39.4	✓	✕ **- Overall: Uncharged, nothing**	✕	N.D.	N.D.	N.D.
T14	β-sheet	N/A	54.8	✓	✕ **- Overall: Uncharged, nothing**	✕	N.D.	N.D.	N.D.
S20	Random coil	N/A	86.9	✓	✕ **- Overall: Uncharged, nothing**	✓ - OH close to P19	N.D.	N.D.	N.D.
T22	Random coil	N/A	46.2	✓	✕ - H-bonding with backbone NH of N25 **- Overall: Uncharged, polar contact**	✕	N.D.	N.D.	N.D.
K27	α-helix	10.53	9.8	✕	✓ - Salt bridge with side chain COOH of D52 - H-bonding with backbone CO of Q41 **- Overall: Charged, (-)**	✓ - NH_2 close to I23	N.D.	N/A	N/A

[a] N/A = not applicable; [b] N-terminal blocking using SNHSA also results in side reaction (acetylation) at Lys and His residues in free peptides, so they are not discussed here.
[c] N.D. = not detected (less than DEPC labeling threshold of 0.01%[5])
* SASA calculated using GETAREA[4], as explained in detail in the Supplemental Experimental Section; ‡ pKa calculated using PROPKA[6]

Ub residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
K29	α-helix	10.40	37.3	✓	✗ - H-bonding with backbone CO of E16 **- Overall: Uncharged, polar contact**	✓ - NH_2 close to V17	N.D.	N/A	N/A
K33	α-helix	10.27	49.4	✓	✗ - H-bonding with backbone CO of T14 **- Overall: Uncharged, polar contact**	✓ - NH_2 close to I13, L15	20 ± 6	N/A	N/A
K48	Random coil	10.31	56.1	✓	✗ - H-bonding with backbone CO of A46 **- Overall: Uncharged, polar contact**	✓ - NH_2 close to A46	11 ± 4	N/A	N/A
T55	Random coil	N/A	32.3	✓	✗ - H-bonding with side chain OH of S57 - H-bonding with backbone NH of D58 **- Overall: Uncharged, polar contact**	✗	N.D.	N.D.	N.D.
S57	α-helix	N/A	59.6	✓	✗ - H-bonding with side chain OH of T55 - H-bonding with backbone CO of P19 **- Overall: Uncharged, polar contact**	✗	N.D.	N.D.	N.D.
Y59	α-helix	9.98	12.9	✗	✗ - H-bonding with backbone NH of E51 **- Overall: Uncharged, polar contact**	✓ - OH close to L50	N.D.	N.D.	N.D.
K63	Random coil	10.66	74.3	✓	✗ **- Overall: Uncharged, nothing**	✗	0.31 ± 0.08	N/A	N/A
S65	β-sheet	N/A	9.2	✗	✗ - H-bonding with backbone CO of Q62 - H-bonding with backbone NH of Q62 **- Overall: Uncharged, polar contact**	✓ - OH close to F45, I61	2.4 ± 0.2	N.D.	N.D.
T66	β-sheet	N/A	40.1	✓	✗ **- Overall: Uncharged, nothing**	✓ - OH close to F4	N.D.	N.D.	N.D.
H68	β-sheet	6.00	49.1	✓	✗ **- Overall: Uncharged, nothing**	✓ - NH close to I44	2.7 ± 0.5	N/A	N/A

Table S3. DEPC modification percentages of nucleophilic residues in hGH intact protein and its proteolytic peptides. Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1HGU. ✓ and ✕ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

hGH residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
T4	Random coil	N/A[a]	100	✓	✕ **- Overall: Uncharged, nothing**	✓ - OH close to P3	N.D.[c]	N.D.	N.D.
S8	α-helix	N/A	40.3	✓	✕ **- Overall: Uncharged, nothing**	✓ - OH close to V186	N.D.	N.D.	N.D.
H19	α-helix	4.78	37.2	✓	✕ **- Overall: Uncharged, nothing**	✓ - imidazole amine close to M15	1.3 ± 0.8 (H19/H22)	N/A[b]	N/A
H22	α-helix	5.16	24.0	✓	✕ **- Overall: Uncharged, nothing**	✓ - imidazole amine close to F26, M171		N/A	N/A
T28	α-helix	N/A	10.7	✕	✓ - H-bonding with side chain COOH of D27 - H-bonding with backbone CO of L24 **- Overall: Charged, (-)**	✓ - OH close to L24	1.3 ± 0.6	N.D.	N.D.
Y29	α-helix	13.59	4.1	✕	✓ - H-bonding with side chain NH_2 of K42 **- Overall: Charged, (+)**	✓ - OH close to Y161	0.2 ± 0.1	N.D.	N.D.
Y36	α-helix	10.35	37.2	✓	✕ **- Overall: Uncharged, nothing**	✓ - OH close to F32	0.4 ± 0.3	N.D.	N.D.
K39	**No information from a crystal structure about this residue**						0.06 ± 0.01	N/A	N/A
K42	Random coil	11.22	36.2	✓	✕ - H-bonding with side chain OH of Y28 **- Overall: Uncharged, polar contact**	✕	0.07 ± 0.02	N/A	N/A
Y43	Random coil	10.05	87.7	✓	✕ **- Overall: Uncharged, nothing**	✕	N.D.	N.D.	N.D.
S44	Random coil	N/A	37.0	✓	✕ - H-bonding with backbone CO of Q41 - H-bonding with backbone NH of F45 **- Overall: Uncharged, polar contact**	✓ - OH close to F45, L46	N.D.	N.D.	N.D.

[a] N/A = not applicable; [b] N-terminal blocking using SNHSA also results in side reaction (acetylation) at Lys and His residues in free peptides, so they are not discussed here.
[c] N.D. = not detected (less than DEPC labeling threshold of 0.01%[5])
* SASA calculated using GETAREA[4], as explained in detail in the Supplemental Experimental Section; ‡ pKa calculated using PROPKA[6]

hGH residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
T51	**No information from a crystal structure about this residue**						N.D.	N.D.	N.D.
S52	Random coil	N/A	65.5	✓	✗ - H-bonding with backbone CO, NH of L53 **- Overall: Uncharged, polar contact**	✗	N.D.	N.D.	N.D.
S56	Random coil	N/A	20.8	✓	✗ - H-bonding with backbone CO of L53 **- Overall: Uncharged, polar contact**	✓ - OH close to L53	0.02 ± 0.01	N.D.	N.D.
S58	Random coil	N/A	3.3	✗	✗ - H-bonding with backbone CO, NH of I59 - H-bonding with backbone CO of F55 **- Overall: Uncharged, polar contact**	✓ - OH close to F177	64 ± 4	N.D.	N.D.
T61	Random coil	N/A	38.7	✓	✓ - H-bonding with side chain NH_2 of K73 **- Overall: Charged, (+)**	✓ - OH close to P62, F177	N.D.	N.D.	N.D.
S63	Random coil	N/A	17.6	✗	✗ - H-bonding with backbone CO of P62 **- Overall: Uncharged, polar contact**	✓ - OH close to F177	N.D.	N.D.	N.D.
T68	**No information from a crystal structure about this residue**						N.D.	N.D.	N.D.
K71	Random coil	10.04	27.1	✓	✗ - H-bonding with backbone CO of Q69 **- Overall: Uncharged, polar contact**	✓ - NH_2 close to I59	N.D.	N/A	N/A
S72	Random coil	N/A	32.2	✓	✗ - H-bonding with backbone NH of L74 **- Overall: Uncharged, polar contact**	✓ - OH close to L74	N.D.	N.D.	N.D.
S80	α-helix	N/A	0.5	✗	✗ - H-bonding with backbone CO of L76 **- Overall: Uncharged, polar contact**	✓ - OH close to L83, V174, L178,	N.D.	N.D.	N.D.
S86	α-helix	N/A	0.1	✗	✗ - H-bonding with backbone CO of Y144 **- Overall: Uncharged, polar contact**	✓ - OH close to F55, W87	N.D.	N.D.	N.D.
S96	Random coil	N/A	49.7	✓	✗ **- Overall: Uncharged, nothing**	✗	N.D.	N.D.	N.D.

hGH residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
S101	Random coil	N/A	33.4	✓	✗ - H-bonding with backbone CO of V103 **- Overall: Uncharged, polar contact**	✗	N.D.	N.D.	N.D.
Y104	Random coil	10.05	82.5	✓	✗ **- Overall: Uncharged, nothing**	✗	N.D.	N.D.	N.D.
S107	Random coil	N/A	26.1	✓	✓ - H-bonding with side chain guanidinium of R20 **- Overall: Charged, (+)**	✓ - OH close to L24	N.D.	N.D.	0.19 ± 0.12
S109	Random coil	N/A	49.3	✓	✓ - H-bonding with sidechain COOH of D106 **- Overall: Charged, (-)**	✗	N.D.	N.D.	N.D.
Y112	α-helix	10.38	29.9	✓	✗ **- Overall: Uncharged, nothing**	✗	N.D.	N.D.	N.D.
K116	α-helix	10.63	40.9	✓	✗ **- Overall: Uncharged, nothing**	✓ - NH_2 close to L88, Y112	N.D.	N/A	N/A
T124	α-helix	N/A	47.2	✓	✗ **- Overall: Uncharged, nothing**	✓ - OH close to L10	N.D.	N.D.	N.D.
S133	Random coil	N/A	80.2	✓	✗ - H-bonding with backbone CO of G132 **- Overall: Uncharged, polar contact**	✗	N.D.	N.D.	N.D.
T136	Random coil	N/A	35.9	✓	✓ - H-bonding with side chain guanidinium of R78 **- Overall: Charged, (+)**	✓ - OH close to L74	0.15 ± 0.09	0.1 ± 0.1	0.08 ± 0.02
K141	Random coil	10.42	90.5	✓	✗ **- Overall: Uncharged, nothing**	✗	18 ± 8	N/A	N/A
T143	Random coil	N/A	96.4	✓	✗ **- Overall: Uncharged, nothing**	✓ - OH close to A145	N.D.	N.D.	N.D.
Y144	Random coil	10.41	31.1	✓	✗ **- Overall: Uncharged, nothing**	✓ - OH close to L82	0.66 ± 0.05	N.D.	0.02 ± 0.03

hGH residue	Secondary structure	pK_a	Solvent accessibility		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
S145	**No information from a crystal structure about this residue**						N.D.	N.D.	N.D.
K146	Random coil	9.4	1.1	✕	✕ **- Overall: Uncharged, nothing**	✓ - NH_2 close to L53, F55	0.09 ± 0.08	N/A	N/A
T149	**No information from a crystal structure about this residue**						N.D.	N.D.	N.D.
S151	Random coil	N/A	56.7	✓	✕ - H-bonding with backbone CO of N150 **- Overall: Uncharged, polar contact**	✕	N.D.	N.D.	N.D.
H152	Random coil	6.38	84.4	✓	✕ **- Overall: Uncharged, nothing**	✕	N.D.	N/A	N/A
K159	α-helix	11.43	70.9	✓	✓ - Salt bridge with sidechain COOH of E34 **- Overall: Charged, (-)**	✕	N.D.	N/A	N/A
Y161	α-helix	11.14	10.0	✕	✕ **- Overall: Uncharged, nothing**	✓ - OH close to L46	N.D.	N.D.	N.D.
T165	**No information from a crystal structure about this residue**						N.D.	N.D.	N.D.
K169	α-helix	10.00	42.8	✓	✕ **- Overall: Uncharged, nothing**	✕	N.D.	N/A	N/A
K173	α-helix	8.71	1.4	✕	✕ - H-bonding with backbone CO of K169 **- Overall: Uncharged, polar contact**	✓ - NH_2 close to F55	5.0 ± 0.4	N/A	N/A
T176	α-helix	N/A	19.0	✕	✓ - H-bonding with sidechain NH_2 of R179 - H-bonding with backbone CO of D172 **- Overall: Charged, (+)**	✕	N.D.	N.D.	N.D.
S185	Random coil	N/A	24.0	✓	✕ - H-bonding with backbone CO of V181 **- Overall: Uncharged, polar contact**	✓ - OH close to F11	N.D.	N.D.	N.D.
S189	Random coil	N/A	96.1	✓	✕ **- Overall: Uncharged, nothing**	✕	N.D.	N.D.	N.D.

SUPPLEMENTAL REFERENCES

1. Verdone, G.; Corazza, A.; Viglino, P.; Pettirossi, F.; Giorgetti, S.; Mangione, P.; Andreola, A.; Stoppini, M.; Bellotti, V.; Esposito, G., The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci.* **2002,** *11* (3), 487-499.
2. Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J., Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **1987,** *194* (3), 531-544.
3. Chantalat, L.; Jones, N. D.; Korber, F.; Navaza, J.; Pavlovsky, A. G., The Crystal Structure of Wild-type Growth Hormone at 2.5 Å Resolution. *Protein Pept. Lett.* **1995,** *2* (2), 333-340.
4. Fraczkiewicz, R.; Braun, W., Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.* **1998,** *19* (3), 319-333.
5. Limpikirati, P.; Hale, J. E.; Hazelbaker, M.; Huang, Y.; Jia, Z.; Yazdani, M.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Covalent labeling and mass spectrometry reveal subtle higher order structural changes for antibody therapeutics. *mAbs* **2019,** *11* (3), 463-476.
6. Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Struct., Funct., Bioinf.* **2005,** *61* (4), 704-721.