Resource Summary Report

Generated by ASWG on Aug 6, 2026

Amino Acid Index Database

RRID:SCR_007044
Type: Tool

Proper Citation

Amino Acid Index Database (RRID:SCR_007044)

Resource Information

URL: http://www.genome.ad.jp/aaindex/

Proper Citation: Amino Acid Index Database (RRID:SCR_007044)

Description: AAindex is a database of numerical indices representing various physicochemical and biochemical properties of amino acids and pairs of amino acids. AAindex consists of three sections now: AAindex1 for the amino acid index of 20 numerical values, AAindex2 for the amino acid mutation matrix and AAindex3 for the statistical protein contact potentials. All data are derived from published literature. An amino acid index is a set of 20 numerical values representing any of the different physicochemical and biological properties of amino acids. The AAindex1 section of the Amino Acid Index Database is a collection of published indices together with the result of cluster analysis using the correlation coefficient as the distance between two indices. This section currently contains 544 indices. Another important feature of amino acids that can be represented numerically is the similarity between amino acids. Thus, a similarity matrix, also called a mutation matrix, is a set of 210 numerical values, 20 diagonal and 20x19/2 off-diagonal elements, used for sequence alignments and similarity searches. The AAindex2 section of the Amino Acid Index Database is a collection of published amino acid mutation matrices together with the result of cluster analysis. This section currently contains 94 matrices. In the release 9.0, we added a collection of published protein pairwise contact potentials to AAindex as AAindex3. This section currently contains 47 contact potential matrices. Sponsors: This work was supported by grants and resources from the Ministry of Education, Culture, Sports, Science and Technology, and the Japan Science and Technology Agency, and the Bioinformatics Center, Institute for Chemical Research, Kyoto University and the Super Computer System, Human Genome Center, Institute of Medical Science, University of Tokyo.

Synonyms: AAindex

Resource Type: data or information resource, database

Defining Citation: PMID:3244698, PMID:9053899, PMID:9847231, PMID:10592278

Keywords: amino acid, biochemical property, mutation, physicochemical property, protein sequence, proteomics, FASEB list

Resource Name: Amino Acid Index Database

Resource ID: SCR_007044

Alternate IDs: nif-0000-02527

Record Creation Time: 20220129T080239+0000

Record Last Update: 20260806T054436+0000

Ratings and Alerts

No rating or validation information has been found for Amino Acid Index Database.

No alerts have been found for Amino Acid Index Database.

Data and Source Information

Source: SciCrunch Registry

Usage and Citation Metrics

We found 205 mentions in open access literature.

Listed below are recent publications. The full list is available at ASWG .

Liang Y, et al. (2025) A deep learning model for prediction of lysine crotonylation sites by fusing multi-features based on multi-head self-attention mechanism. Scientific reports, 15(1), 18940.

Charoenkwan P, et al. (2025) PSR-MAPMS: A new approach for the interpretable prediction of myelin autoantigenic peptides in multiple sclerosis using multi-source propensity scores. Protein science : a publication of the Protein Society, 34(8), e70010.

Zhang Q, et al. (2025) Machine learning-guided evolution of pyrrolysyl-tRNA synthetase for improved incorporation efficiency of diverse noncanonical amino acids. Nature communications, 16(1), 6648.

Li R, et al. (2025) AOPxSVM: A Support Vector Machine for Identifying Antioxidant Peptides Using a Block Substitution Matrix and Amino Acid Composition, Transformation, and Distribution Embeddings. Foods (Basel, Switzerland), 14(12).

Cun Y, et al. (2025) SITA: Predicting site-specific immunogenicity for therapeutic antibodies. Journal of pharmaceutical analysis, 15(6), 101316.

Gomez-Uribe CA, et al. (2025) Designing diverse and high-performance proteins with a large language model in the loop. PLoS computational biology, 21(6), e1013119.

Maiti S, et al. (2025) Proteome-wide computational analyses reveal links between protein condensate formation and RNA biology. Science advances, 11(49), eady1420.

Liu T, et al. (2025) Effective Gene Expression Prediction and Optimization from Protein Sequences. Advanced science (Weinheim, Baden-Wurttemberg, Germany), 12(8), e2407664.

Elreify HM, et al. (2025) An efficient machine-learning framework for predicting protein post-translational modification sites. Scientific reports, 15(1), 31179.

Keresztes L, et al. (2025) Navigating Homogeneous Graph Paths Through Amyloidogenic and Non-Amyloidogenic Hexapeptides. Journal of computational chemistry, 46(26), e70238.

Pavesi A, et al. (2025) Covariation of Amino Acid Substitutions in the HIV-1 Envelope Glycoprotein gp120 and the Antisense Protein ASP Associated with Coreceptor Usage. Viruses, 17(3).

G??ktepe YE, et al. (2025) Protein-protein interaction prediction using enhanced features with spaced conjoint triad and amino acid pairwise distance. PeerJ. Computer science, 11, e2748.

Landwehr GM, et al. (2025) Accelerated enzyme engineering by machine-learning guided cell-free expression. Nature communications, 16(1), 865.

Li M, et al. (2025) MVSF-AB: accurate antibody-antigen binding affinity prediction via multi-view sequence feature learning. Bioinformatics (Oxford, England), 41(5).

Duhan N, et al. (2025) PRGminer: harnessing deep learning for the prediction of resistance genes involved in plant defense mechanisms. Frontiers in plant science, 16, 1411525.

Liu S, et al. (2025) FusionEncoder: identification of intrinsically disordered regions based on multi-feature fusion. Bioinformatics (Oxford, England), 41(7).

Harun-Or-Roshid M, et al. (2025) A genetic algorithm-based ensemble model for efficiently identifying interleukin 6 inducing peptides. Scientific reports, 15(1), 21213.

Li X, et al. (2025) Predicting nucleic acid binding sites by attention map-guided graph convolutional network with protein language embeddings and physicochemical information. Briefings in bioinformatics, 26(5).

Duan Z, et al. (2025) Artificial intelligence-driven framework for discovering synthetic binding protein-like scaffolds from the entire protein universe. Briefings in bioinformatics, 26(5).

Silva A, et al. (2025) Improving protein interaction prediction in GenPPi: a novel interaction sampling approach preserving network topology. BMC bioinformatics, 26(1), 296.