Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d6sc01551d

All publication charges for this article have been paid for by the Royal Society of Chemistry

PpF: a density functional fine-tuned for noncovalent interactions of protein and peptide residues

Yini Zhou,[ab] Tao Li,[ab] Yaqi Li,[ab] Jianda Yue,[ab] Qifeng Tian,[ab] Zhonghua Liu,[ab] Donald G. Truhlar *[c] and Ying Wang *[ab]

The essence of protein–peptide interactions lies in the noncovalent interactions between amino-acid pairs; accurately calculating interaction energies of these pairs is crucial for modelling the tertiary structure and protein–peptide interactions. However, available functionals for density functional theory have insufficient accuracy for many noncovalent interactions. To address this challenge, we developed a new functional called PpF by starting with the broadly trained CF22D model as a foundation model and fine tuning it by using it with more specific data on interactions of capped amino acids. The PpF model is specifically designed for noncovalent interactions of pairs of amino acids. First, based on the LEADS-PEP dataset for assessing peptide docking performance, we constructed the amino-acid pair structures dataset (called AAPS260) containing 260 pairs of noncovalently interacting capped amino acids, from which we selected 36 representatives. We performed DLPNO-CCSD(T) calculations on these pairs to determine reference energies for a new training dataset with 12 interaction energies and a new testing dataset with 24 interaction energies. We used an iterative supervised training strategy to optimize parameters for an exchange–correlation functional with a damped dispersion term; the loss function involves 89 previously defined datasets augmented by the new training dataset (AAIE12) with a performance-triggered determination of its weight. This produces the PpF functional. We find that the PpF functional outperforms other functionals on the training set (AAIE12), test set (AAIE24) and the Side Chain Atlas of Interactions (SCAI) dataset. It also does very well on the JSCH, GMTKN55, and MGCDB84_NC databases. The PpF functional is then used to establish the Amino-Acid Interaction Energy benchmark dataset, which is called AAIE260. This work produces a new density functional, a structural dataset of pairs of capped amino acids, a benchmark dataset of the interaction energies of these pairs, and a reliable computational method for exploring protein–peptide binding mechanisms.

Received 24th February 2026
Accepted 28th May 2026

DOI: 10.1039/d6sc01551d

rsc.li/chemical-science

1. Introduction

Protein–peptide interactions are fundamental features of molecular recognition and activity regulation in living organisms; they play an essential role in key biological processes such as signal transduction, gene expression, cell-cycle control, metabolic regulation, and immune responses.[1–3] These molecular interaction networks promote the precise transmission of signals both inside and outside cells, and they maintain cellular and organismal homeostasis through dynamic regulation of protein conformational changes and functional states.[4,5] At a molecular level, the binding between proteins and peptides involves a complex network of noncovalent interactions between amino acid residues.[6,7] Accurately quantifying the interaction energies, stability, and specificity that drive protein–peptide interactions can effectively identify the core residues and key interaction patterns that drive these processes.[8–10] Quantitative analysis of the interactions not only allows for a more comprehensive and detailed understanding of the binding mechanisms between proteins and peptides but also provides a theoretical foundation and practical guidance for the selection of key binding sites, optimization of peptide sequences, and prediction of binding affinities in drug design.

Density functional theory (DFT),[11] in particular Kohn–Sham density functional theory (KS-DFT),[12–15] is widely regarded as the most effective method for balancing predictive accuracy and computational efficiency when using electronic structure methods for modelling main-group and transition-metal chemistry. However, no functional has yet achieved true

[a]The National and Local Joint Engineering Laboratory of Animal Peptide Drug Development, College of Life Sciences, Hunan Normal University, Changsha 410006, Hunan, China

[b]Peptide and Small Molecule Drug R&D Plateform, Furong Laboratory, Hunan Normal University, Changsha, Hunan, 410081, China

[c]Department of Chemistry, Chemical Theory Center, and Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455-0431, USA. E-mail: truhlar@umn.edu; wangyin@hunnu.edu.cn

universality. For example, the ωB97X-D functional[16] performs well for thermochemistry, kinetics, and noncovalent interactions, but its performance in predicting the proton affinity of amino acids is less satisfactory.[17] For another example, the M06-2X functional[18] performs well for most main-group conformational analysis, thermochemistry, kinetics, and noncovalent interactions, but its performance in predicting the conformational energies of amino acids and peptides is less satisfactory.[19]

The noncovalent network of amino-acid interactions in proteins is highly intricate, involving hydrogen bonds, electrostatic interactions, van der Waals forces, long-range and medium-range dispersion effects, cooperative interactions, hydrophobic effects, and local structural changes.[20] It is challenging for density functionals to accurately capture all these aspects. To address this issue, the present study attempted to improve the accuracy of calculations of amino-acid pair interactions by developing a density functional specifically designed for this kind of interaction.

This first challenge in this work was that we lacked a high-quality benchmark dataset to guide the development. To remedy this, we created a benchmark dataset, and here we make it available for development and validation of next-generation computational methods.

In previous work, we developed a density functional named CF22D (Chemistry Functional 2022 with damped Dispersion)[21] that is designed for application across diverse chemical fields. CF22D was developed by utilizing physical descriptors, extensive databases, and supervised learning. In addition to orthodox exchange–correlation terms, CF22D includes a molecular mechanical damped dispersion term, and it achieves broad chemical accuracy for multiple targets, including reaction barriers, isomerization energies, thermodynamic properties, and weak interactions. The present study uses CF22D as a foundation model and uses continued learning to further optimize and develop a functional with enhanced performance for amino-acid pair interactions in protein–peptide systems; the resulting functional is called protein–peptide functional (PpF).

Previous studies have shown that developing specialty functionals based on high-accuracy general functionals is feasible. For example, the M06CR functional[22] (for reactions of Criegee intermediates) and the M06-HS functional[23] (for hydrogen transfer reactions of peroxy radicals) have shown significant improvements in computational accuracy for their specific targets. This approach is an extension of the specific reaction parameter (SRP)[24–26] method. In the SRP approach, one trains an electronic structure method for a single reaction or small range of related reactions or interactions. In the extended SRP approach employed here, one adds specific reaction data or specific-range data to a general-purpose model to improve the applicability of the general-purpose model in a particular direction.

In the context of machine learning, one can say that the CF22D functional is serving as a broadly applicable foundation model that was trained on broad data and the PpF functional is the result of fine tuning it to a specific purpose by leveraging the pretrained knowledge.

2. Materials and methods

2.1. Data collection and preprocessing

2.1.1. Construction of the AAPS260 dataset. We start with the LEADS-PEP dataset[27] published by Windshügel and coworkers in 2016. This dataset provides a collection of protein–peptide complexes, and it is widely used for studying of protein–peptide interactions and for benchmarking of peptide docking software.[28–30] The dataset consists of 53 protein–peptide complexes sourced from the Protein Data Bank, with peptide fragments containing between 3 and 12 amino-acid residues.

As illustrated in Fig. 1, we first extracted all closely interacting residue pairs from each protein–peptide complex structure. Specifically, for each residue on the peptide segment, we identified all protein residues within a 3 Å distance. Each such "peptide residue–protein residue" combination was defined as an initial interacting pair. Applying this selection procedure to the 53 complexes yielded a total of 506 initial residue pairs. To precisely maintain the local conformation of each residue after removal from the original protein and peptide chain environment, we performed the Molecular Fractionation with Conjugate Caps (MFCC)[31] treatment on each of the residues of each of the pairs. The core of this technique lies in capping the target residue upon cleavage from its backbone by retaining and utilizing specific atoms from its adjacent residues as "caps". For internal residues within the peptide chain, we retained the C, O and CA atoms of the preceding residue to form an acetyl (ACE) cap, connecting to the N-terminus of the current residue; simultaneously, we retained the N and CA atoms of the following residue to form an *N*-methyl amide (NME) cap, connecting to the C-terminus of the current residue. This process generates stand-alone, chemically complete fragment with the structure "ACE-target residue-NME" (refer to Fig. 1C for illustration). For N-terminal residues, the original terminal state was retained along with the N and CA atoms of the next residue for NME capping; for C-terminal residues, the inherent terminal oxygen atom (OXT) of the original terminus and the C, O and CA atoms of the preceding residue were retained for ACE capping.

Subsequently, hydrogen atoms were added to all amino-acid residues according to the parameters of the ff19SB[32] force field, and structural optimization was performed on the added hydrogen atoms. Then, to avoid structural redundancy, a structural similarity assessment was conducted for the amino-acid residue pairs, and residue pairs with a root mean square deviation (RMSD) of less than 2 Å were removed, resulting in 260 amino-acid pairs. This forms the amino-acid pair structures dataset, called AAPS260.

Finally, we assigned uniform and physiologically relevant protonation states and atomic charges to all systems within the AAPS260 dataset. The charge states for all residues were set with reference to physiological pH conditions (7.4): acidic residues (ASP, GLU) carried a −1 charge on their side chains, and basic residues (LYS, ARG) carried a +1 charge on their side chains. Because the imidazole ring of histidine can be protonated at two possible positions, we distinguish the HID tautomer (protonated at the δ nitrogen, 6 cases) from the HIE tautomer

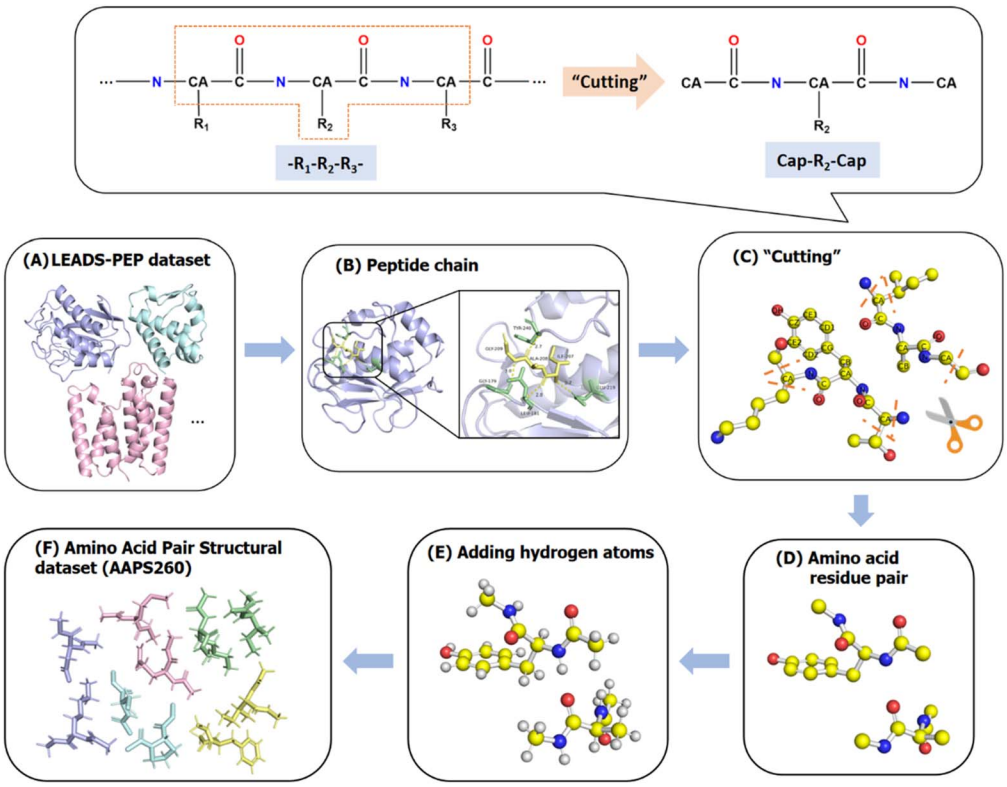

Fig. 1 Construction process of the AAPS260 dataset.

(protonated at the ε nitrogen, 6 cases). Furthermore, to simulate the ionization state of the original peptide chain terminal residues, if a residue was at the N-terminus in the original peptide chain, the net charge of its entire capped fragment was set to +1; if it was at the C-terminus, the net charge was set to −1.

2.1.2. Construction of the AAIE12 and AAIE24 datasets. We selected 36 representative pairs of capped amino acids from the AAPS260 dataset and calculated their energies using a high-level wave function method specified below. We also calculated the energies of the separate capped amino acids (with all atoms, including the hydrogens, frozen in the geometry they had in the pair), and we calculated the interaction energies as the energy of the pair minus the sum of the energies of the two separated species. We randomly selected 12 systems for training, and their interaction energies are denoted AAIE12; the remaining 24 systems were used for testing, and their interaction energies are denoted as AAIE24. Structural models of the 36 pairs of capped amino acids are shown in Fig. 2.

The CCSD(T)/CBS method is often the method of choice for accurate calculations of noncovalent interactions.[33–35] However, due to its high computational cost, it is challenging to apply to larger molecular systems. Therefore, we chose the Domain-based Local Pair Natural Orbital Coupled Cluster (DLPNO-CCSD(T)) method[36] to calculate the reference values. These calculations used the def2-QZVPPD[37] basis set and its corresponding def2-QZVPPD/C auxiliary basis set. The DLPNO parameters were set to the TightPNO setting (TCutPairs = 10^{-5}, TCutPNO = 10^{-7}, and TCutMKN = 10^{-3}) and the TightSCF setting (energy change = 10^{-8} au) to promote the convergence and reliability of the calculations.

2.1.3. Construction of the AAIE260 dataset. To create a larger dataset for future use, we employed the PpF density functional with the ma-TZVP basis set[38] to calculate interaction energies of all 260 pairs of capped amino acids in the AAPS260 dataset. The hydrogen atom positions for these calculations were optimized with the B3LYP-D3(BJ)[39] density functional and the 6-311G** basis set.[40] The interaction energies of all 260 pairs calculated this way constitute the Amino-Acid Interaction Energy dataset, denoted AAIE260.

2.2. Optimization of the density functional

2.2.1. The functional form. The functional form of the PpF exchange-correlation energy is the same as that of CF22D functional,[21] and is given by

$$E^{\mathrm{PpF}} = \frac{X}{100}E_X^{\mathrm{HF}} + E_{\mathrm{nXC}} + E_{\mathrm{C}} + E_{\mathrm{disp}} \tag{1}$$

X is the percentage of HF exchange E_X^{HF}, E_{nXC} is a nonseparable exchange-correlation term,

$$E_{\mathrm{nXC}} = \int \mathrm{d}r \sum_{\sigma=\alpha}^{\beta} \rho_\sigma \left\{ \begin{array}{l} \varepsilon_{x\sigma}{}^{\mathrm{LSDA}}(\rho_\sigma) \sum_{i=0}^{3} \sum_{j=0}^{3-i} \sum_{k=0}^{5-i-j} a_{ijk} \cdot \\ \{v_{x\sigma}(\rho_\sigma)\}^i \{u_{x\sigma}(s_\sigma)\}^i \{w_\sigma(\rho_\sigma, \tau_\sigma)\}^k \end{array} \right\} \tag{2}$$

E_{C} is a dynamic correlation term,

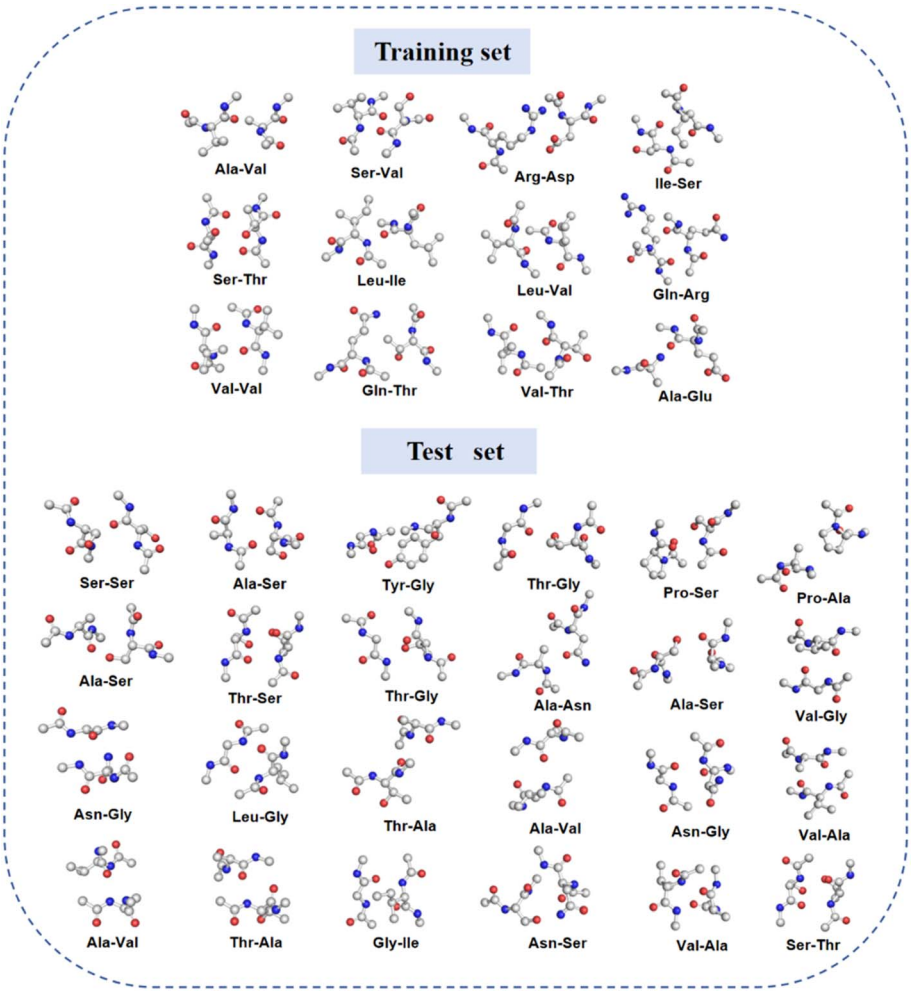

Fig. 2 Structures of the pairs of capped amino acids in the training and test sets.

$$E_C = \int d\mathbf{r}\rho\varepsilon_C^{LSDA}(\rho_\alpha, \rho_\beta)\left(\sum_{i=0}^{8} b_i\{w(\rho,\tau)\}^i\right) + \int d\mathbf{r}\rho H^{PBE}(\rho_\alpha, \rho_\beta, s)\left(\sum_{i=0}^{8} c_i\{w(\rho,\tau)\}^i\right) \quad (3)$$

$$s_\sigma = \frac{|\nabla\rho_\sigma|}{\rho\sigma^{4/3}} \quad (4)$$

In eqn (2) and (3), ρ_α and ρ_β are the up-spin and down-spin electron densities at the spatial point r, ρ is their sum, τ_α and τ_β are the spin-up and spin-down kinetic energy density. The functions $v_{X\sigma}$, $u_{X\sigma}$, w_σ, $\varepsilon_{X\sigma}^{LSDA}$, ε_C^{LSDA} and H^{PBE} are the same as used in the CF22D functional.

E_{disp} is a molecular mechanics damped-dispersion term,

$$E_{disp} = -\frac{1}{2}\sum_{AB}\frac{C_6{}^{AB}}{r_{AB}{}^6} f_{d,6}(r_{AB}) \quad (5)$$

$$f_{d,6}(r_{AB}) = \frac{1}{1 + 6\left(r_{AB}/s_{r,6}R_0{}^{AB}\right)^{-12}} \quad (6)$$

The scaling parameter $s_{r,6}$ is set to 1.53 Å, consistent with CF22D. $R_0{}^{AB}$ is the D3(0) dispersion coefficient. The PpF parameterization optimizes the 59 linear coefficients of eqn (1)–(3), that is, X, a_{ijk}, b_i and c_i. Each linear parameter multiplies an integral, and these integrals are called features. The $R_0{}^{AB}$ parameters are unchanged from their original[41] values.

The parameterization minimizes a loss function defined by

$$L = \sum_{n=1}^{K} R_n/I_n + \lambda S \quad (7)$$

$$S = \sum_{i=0}^{3}\sum_{j=0}^{3-i}\sum_{k=0}^{4-i-j}\left(a_{ijk} - a_{ijk+1}\right)^2 + \sum_{i=0}^{7}\left(b_i - b_{i+1}\right)^2 + \sum_{i=0}^{7}\left(c_i - c_{i+1}\right)^2 \quad (8)$$

where K is the number of training datasets, totaling 90, which includes 89 datasets selected from DDB22 and the AAIE12 dataset developed in this study, R_n is the root-mean-square error of dataset n, I_n is the inverse weight assigned to dataset n, S is a regularization term that serves as a smoothness restraint, and λ is the smoothing coefficient which is set equal to 0.01 for the

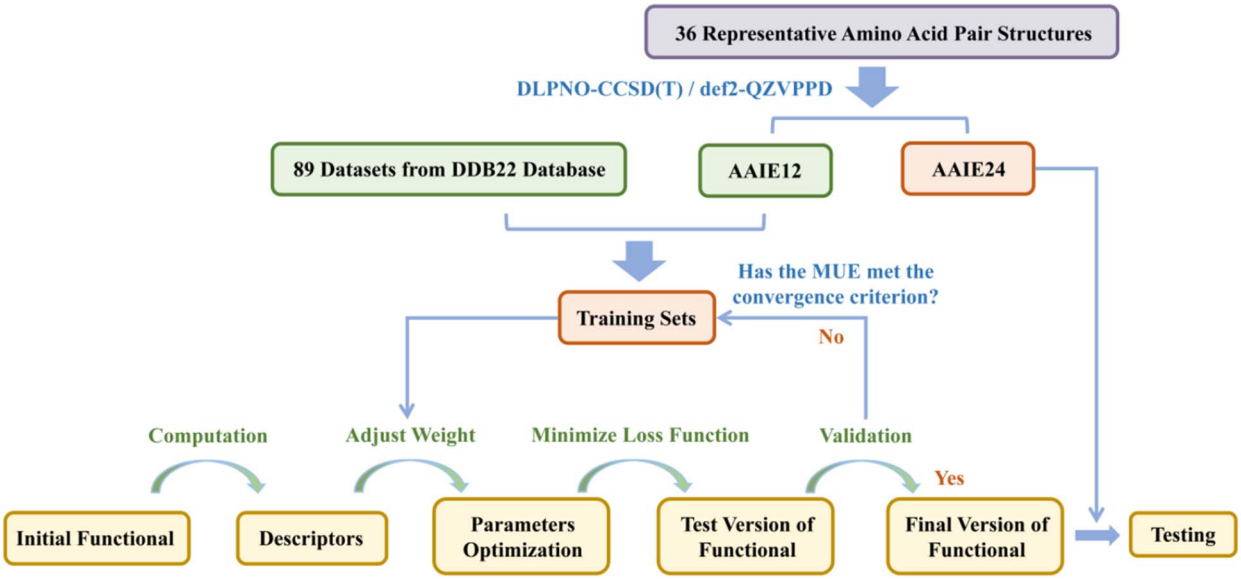

Fig. 3 Training process for the PpF functional.

PpF functional. The final parameters of the optimized PpF functional are provided in Supplementary Table S1. The initial inverse weights are those used for training CF22D, and we modified the inverse weights of AAIE12 dataset based on the convergence results. The inverse weight of each dataset used for developing the final version of the PpF functional is listed in SI Table S2.

2.2.2. Parameters optimization. Optimization of the new density functional was accomplished with an iterative supervised learning method that that builds on the optimization used[21] in the development of CF22D. The process is illustrated in Fig. 3, and the steps are as follows:

2.2.2.1 Create PpF training database. The database constructed for training the PpF model consists of 90 datasets. It is primarily derived from the DDB22 database, specifically selecting the 89 datasets (comprising 1886 data) that were utilized in the initial training phase of CF22D. This was supplemented by 12 data from the AAIE12 dataset. In total, 1898 data were employed in this study's training.

2.2.2.2 Calculate the initial electronic densities. The initial electron densities of all systems in the training set were calculated by using the CF22D functional.

2.2.2.3 Generate initial descriptors. Based on the electron densities computed by the functional from the previous step, the descriptors in the PpF functional described by eqn (1) were calculated for all systems in the training set.

2.2.2.4 Parametrization of the functional. The goal was not to find the global minimum of a predefined loss function but rather to improve performance for the specific target while maintaining good broad accuracy. This is done by adjusting the inverse weights in the loss function and optimizing the functional for various sets of inverse weights. In the initial stage, we attempted to fine-tune the weights for both the AAIE12 dataset and the other datasets simultaneously. However, we found that optimizing only the inverse weight of the AAIE12 dataset led to better results. Therefore, we decided to fine-tune only the inverse weight for the AAIE12 dataset, while keeping the inverse weights for the other datasets the same as used in optimizing CF22D. The inverse weights of each of the training datasets are provided in SI Table S2.

Having chosen to vary only one inverse weight, the next issue is optimizing that inverse weight and the 59 linear coefficients in the functional to minimize the loss function for a given set of inverse weights. We caried out such optimizations by an iterative process. For a given inverse weight, the energies of all systems in the training set were calculated, along with the mean unsigned error (MUE) for each dataset within it. When the MUE of the AAIE12 dataset was higher than that of other functionals, we decreased the inverse weight of the AAIE12 dataset and repeated the process. We conducted multiple optimization runs with different initial conditions and compared the results to ensure consistency and reliability in the convergence of the iterative process. Through multiple optimization trials, we confirmed that each optimization with a given set of inverse weights consistently converged to the same solution, demonstrating the stability of the solution.

2.2.2.5 Finalize. We rounded all linear parameters except X in eqn (1) to nine decimal places, and we rounded X to six significant figures. These are the final parameters of the functional. We used these to recompute the densities and descriptors and to calculate all final published PpF results, including the final MUEs.

2.2.3. Databases. To evaluate the performance of the PpF functional, we compared the results obtained using PpF with those from several mainstream functionals across the amino-acid pair dataset we developed, as well as several literature datasets. These tests include AAIE12, AAIE24, Side Chain Atlas of Interactions (SCAI),[42] the Jurečka–Šponer–Černý–Hobza database of interaction energies of small model complexes, DNA base pairs, and amino acid pairs (JSCH),[43] three

noncovalent interaction databases (NCED, NCEC, NCD) from the MGCDB84 database,[44] and the GMTKN55 database.[45]

2.3. Computational details

The energy calculations in this study were performed using the following programs: ORCA 5.0.3 (ref. 46) for DLPNO-CCSD(T) ωB97M-V, ωB97X-V, r2SCAN and r2SCAN-3c functionals, Gaussian 09 (ref. 47) with the Minnesota Gaussian Functional Module (MN-GFM) 6.11 (ref. 48) for the revM11, revM06, revM06-L and M06-SX functionals, a locally modified version of Gaussian 16 (ref. 49) for the PpF and CF22D functionals, Gaussian 16 for the other functionals, and PSI4 (ref. 50) for interaction energy decomposition calculations carried out by the sSAPT0 method with the jun-cc-pVDZ basis set.

All functional results for the training set (AAIE12), test set (AAIE24), and the external test sets SCAI and JSCH were obtained from calculations performed in this work. For the MGCDB84_NC and GMTKN55 databases, only the results for the PpF, r2SCAN, and r2SCAN-3c functionals were obtained in the present work, whereas the other results used for comparison were taken from the literature.[21]

The choice of basis set depends on the database. We used ma-TZVP for AAIE260 and SCAI, aug-cc-pVDZ for JSCH, def2-QZVPPD for MGCDB84_NC, and def2-QZVP for GMTKN55.

2.4. Evaluation metric

The MUE is used as the evaluation metric for assessing the performance of the PpF functional across the 90 datasets in this study. The MUE measures the accuracy of the method by calculating the average of the absolute errors. The formula for calculating the MUE of a dataset is as follows:

$$\mathrm{MUE} = \frac{1}{N}\sum_{i=1}^{N}\left|E_i^{\mathrm{ref}} - E_i^{\mathrm{calc}}\right| \tag{9}$$

where E_i^{ref} is the i-th reference value, E_i^{calc} is the i-th computed value, and N is the amount of data in the dataset.

3. Results and discussion

By comparing the MUEs of various functionals across these datasets, we assessed the accuracy and robustness of the PpF functional for different types of interactions. SI Table S3 shows all the functionals tested. The detailed results for all datasets are summarized in the SI.

3.1. AAIE12 and AAIE24

3.1.1. Overall performance. The results of the PpF functional were compared with those of 27 mainstream functionals. The interaction energies for these systems and the detailed results for all functionals are listed in SI Tables S4 to S6. Fig. 4A shows the result for the 15 best performing functionals.

We find that the PpF functional achieves the best overall performance on both datasets. The MUEs for the training set (AAIE12) and test set (AAIE24) are 0.11 and 0.10 kcal mol^{-1}, respectively, which are 1.33 and 1.22 kcal mol^{-1} lower than the average MUEs of all functionals tested. The ωB97X-V functional yields MUEs of 0.14 and 0.17 kcal mol^{-1}, respectively, and the PBE0-D3(BJ) functional yields 0.15 and 0.22 kcal mol^{-1}; these are the second and third best performances. The popular B3LYP-D3(BJ) functional has MUEs of 0.24 kcal mol^{-1} for both the training and test sets, ranking sixth.

The CF22D functional yields MUEs of 0.60 and 0.40 kcal mol^{-1} for the training and test sets, respectively, ranking eighth among the 28 tested functionals; this shows its limitations in accurately describing amino-acid pair interactions. The PpF outperforms CF22D significantly for both the training and test sets, showing that the present extended SRP strategy of fine tuning CF22D is successful.

Noncovalent interactions typically rely on the accurate description of medium-range electronic correlation effects and dispersion forces. PBE0, PBE, B3LYP, and TPSS do not include explicit dispersion terms and have lower accuracy for these interactions. These functionals ranked last in both the training and test sets. However, after introducing the post-self-consistent-field D3(BJ) damped dispersion terms, their performance for noncovalent interactions. Improved significantly. For instance, PBE0-D3(BJ) ranked third in both the training and test sets (0.15 kcal mol^{-1} and 0.22 kcal mol^{-1}), showing notable improvement.

Some functionals, specifically PpF, ωB97X-D, B3LYP-D3(BJ), M06-2X-D3(0), and PW6B95-D3(0), in Fig. 4A show only small MUE differences between the training and test sets, indicating their good stability and transferability in describing amino-acid-pair noncovalent interactions. This suggests that these functionals provide a relatively balanced description of the dominant interaction components in the present systems, including electrostatic interactions, hydrogen bonding, and dispersion interactions. As a result, even when the training and test sets differ in specific geometries and interaction patterns, their overall errors remain stable. By contrast, the lower-ranked functionals exhibit more pronounced error differences between the training and test sets. For example, although PBE-D3(BJ), M06L-D3(0), and TPSS-D3(BJ) have been retooled by the inclusion of damped dispersion terms, their treatment of exchange, polarization, and medium-range correlation effects still appears to be less balanced than that of the top-performing functionals, so their errors fluctuate more noticeably when the test set contains interaction patterns that differ somewhat from those in the training set. In addition, functionals such as revM06, TPSS-D3(BJ), and M11 already show relatively large errors on the training set, suggesting a weaker overall suitability for this type of amino-acid pair noncovalent interaction, so even small compositional differences between the training and test sets can lead to more pronounced error differences.

We compared the computational cost of PpF to that of other commonly used functionals using the ALA–VAL residue pair system (50 atoms) from the test set, based on single-core CPU time measured on a dual-socket AMD EPYC 7452 server (2 sockets, 32 cores per socket, 128 logical threads in total, x86_64 architecture). Full timings are provided in SI Table S7. We compare the computer times as we found hem even though it is well known that timings have some variation depending on the

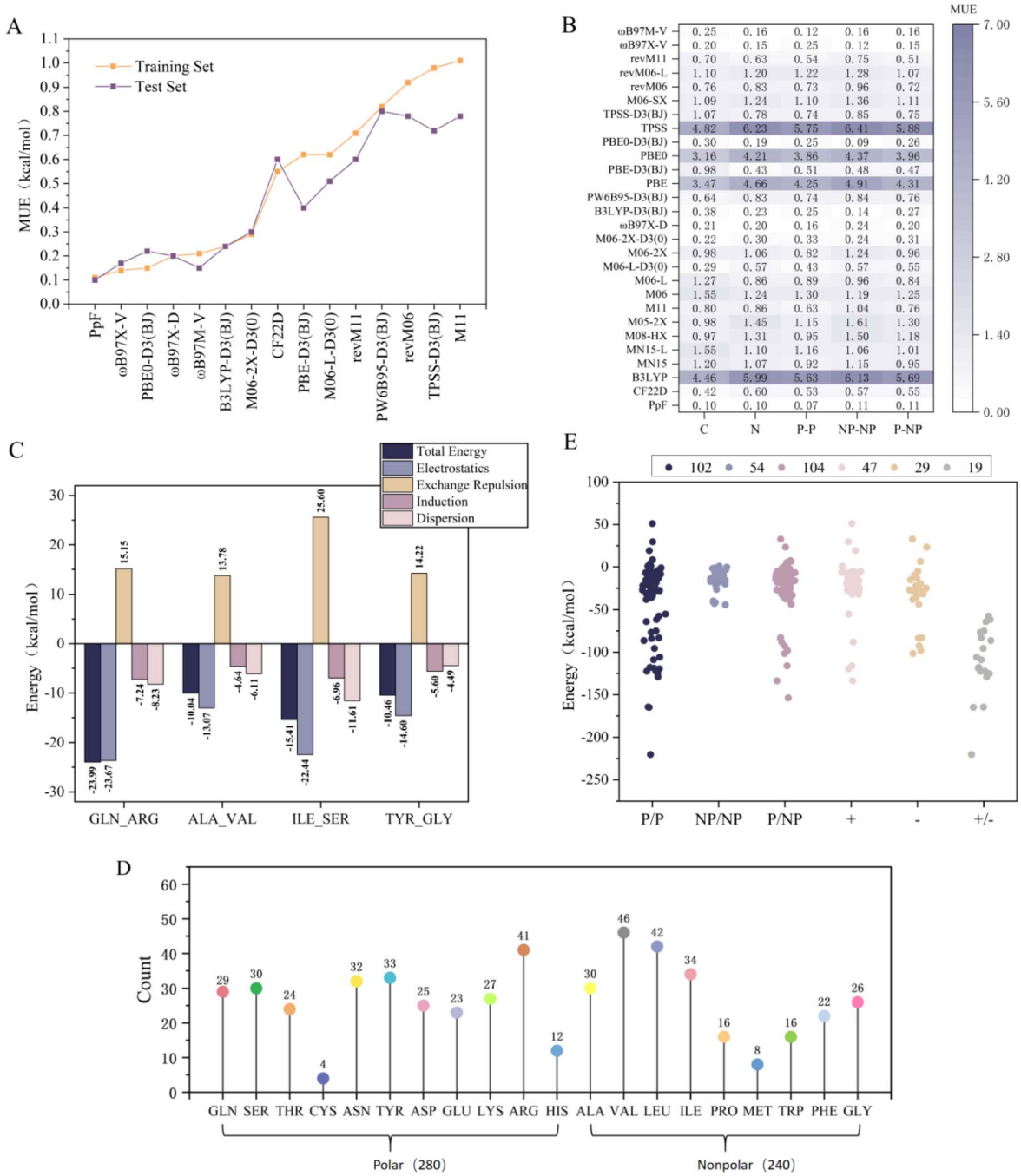

Fig. 4 Results for datasets created in the present study. (A) MUEs (kcal mol^{-1}) of the 15 best performing functionals on the amino-acid pair training and test sets. (B) MUESs (kcal mol^{-1}) of 28 functionals in amino-acid pair systems with different polarity and charge combinations, displayed as a heatmap. Darker colors represent larger MUE values. "C" represents combinations containing charged amino acids, "N" represents combinations with only neutral amino acids, "P–P" represents polar–polar combinations, "NP–NP" represents nonpolar–nonpolar combinations, and "P–NP" represents polar-nonpolar mixed combinations. (C) Interaction energy decomposition of four amino-acid pair systems (kcal mol^{-1}). (D) Amino-acid composition in the AATE260 dataset. (E) Distributions ofi nteraction energies (kcal mol^{-1}) in the AATE260 dataset. The 260 amino-acid pairs in the AATE260 dataset were classifiedi nto the following categories: polar–polar (P/P, 102 pairs), nonpolar–nonpolar (NP/NP, 54 pairs), and polar-nonpolar (P/NP, 104 pairs). Among the pairs containing one or two polar amino acids, further subdivisions were made: combinations containing one or two positively charged amino acids (+, 47 pairs), combinations containing one or two negatively charged amino acids (–, 29 pairs), and mixed combinations containing both positively and negatively charged amino-acids (+/–, 19 pairs).

implementation, software, hardware, and run-time parameters. The results show that the computational time for PpF is 6.9 h, which is slightly longer than that for B3LYP (5.7 h) and PBE0-D3(BJ) (5.8 h), but significantly lower than that for ωB97X-D (8.4 h) and M11 (10.1 h). PpF, was derived from CF22D by fine-tuning, and in our test, it required slightly more computer

time (6.9 h *vs.* 6.3 h), but since it provides a more accurate description of complex noncovalent interactions in amino acid pairs, this small increase in computational cost is justified.

3.1.2. Performance for different classes of amino-acid interactions. The 36 pairs of capped amino acids were divided into three groups: polar–polar (P–P, 8 pairs), nonpolar–nonpolar (NP-NP, 11 pairs), and polar–nonpolar mixed (P–NP, 17 pairs). The pairs were alternatively divided into charged (C, 3 pairs) and neutral (N, 33 pairs) pairs. The performance of 28 functionals for these different classes of interactions was then analyzed. Fig. 4B shows the MUEs of the 28 functionals for different classes of interactions.

PpF shows the broadest performance among all the tested functionals, consistently exhibiting a relatively low MUE of ≤0.11 kcal mol^{-1} across different classes of interactions. For polar–polar interactions, the MUE of PpF is 0.07 kcal mol^{-1}, significantly lower than that of other functionals. Next are ωB97X-V, PBE0-D3(BJ), ωB97X-D, and ωB97M-V, which have MUEs of ≤0.30 kcal mol^{-1} for all classes of interactions. In contrast, PBE0, PBE, B3LYP, and TPSS performed poorly, with MUEs ranging from 3.16 kcal mol^{-1} to 6.41 kcal mol^{-1} across different interaction classes.

Overall, the PpF functional ranks first for four of the classes (C, N, P–P, and P–NP) and second for the NP–NP class, for which it is surpassed only by PBE0-D3(BJ). This demonstrates the broad accuracy of PpF for noncovalent interactions of amino acids.

3.1.3. Interaction energy decomposition. We selected four representative pairs, namely GLN–ARG, ALA–VAL, ILE–SER and TYR–GLY, for energy decomposition analysis by the sSAPT0/jun-cc-pVDZ method, and the results shown in Fig. 4C. This shows that the interaction energy is mainly governed by electrostatic interactions and exchange repulsion, followed by induction, with (damped) dispersion contributing the least.

For the GLN–ARG and TYR–GLY interactions, the contribution of electrostatics exceeds that of exchange repulsion, whereas for the ALA–VAL and ILE–SER interactions, the opposite is observed. This difference primarily arises from the charge differences and polarity differences of the amino-acid side chains. In the GLN–ARG and TYR–GLY pairs, the amino acids contain highly polar or charged groups such as the amide group (–$CONH_2$) of GLN, the guanidinium group (–$C(NH_2)_2^+$) of ARG, and the phenolic hydroxyl group (–OH) of TYR. These groups promote strong hydrogen bonds and charge–dipole interactions, leading to a significant increase in electrostatic attraction, while the relatively large intermolecular distance and limited orbital overlap result in a smaller exchange repulsion.

In contrast, the ALA–VAL and ILE–SER pairs have nonpolar or weakly polar residues. The former represents a typical hydrophobic nonpolar–nonpolar pair, while the latter includes only a weakly polar hydroxyl group. These pairs lack strong charge or dipole interactions and are instead stabilized primarily by van der Waals and electron cloud contact interactions. Because of their shorter intermolecular distances, electron cloud overlap is greater, thereby increasing Pauli (exchange) repulsion energy, which makes exchange repulsion slightly stronger than electrostatic attraction. Overall, polar or charged amino-acid pairs are dominated by electrostatic attraction, whereas nonpolar or weakly polar amino-acid pairs are governed mainly by short-range exchange repulsion and dispersion interactions, reflecting the fundamental differences in the nature of their intermolecular interactions.

In most systems, the contribution of induction to overall stability is minimal, typically representing the smallest component of the interaction energy. However, in the TYR–GLY system, the dispersion term is the least significant. This deviation likely arises because the phenolic hydroxyl group of TYR forms a stable hydrogen-bond network with the backbone of GLY, thereby enhancing the electrostatic and induction contributions.

3.2. AAIE260 dataset

The newly developed PpF functional for amino-acid pair interactions was used to calculate the interaction energies for all pairs of capped amino acids in the AAPS260 dataset. These interaction energies constitute the AAIE260 reference dataset, which is provided in SI Table S8.

A statistical analysis of the 520 capped amino acid structures in the AAIE260 dataset was conducted according to their classification according to polar and nonpolar side chains (Fig. 4D). Among these, 240 are nonpolar, accounting for 46%, and 280 are polar, accounting for 54%.

As shown in Fig. 4E, there are significant differences in interaction energies across different polarity and charge combinations. Lower signed interaction energies correspond to stronger binding strengths. As compared to NP–NP, the energy distribution for P–P and P–NP pairs is wider, and the overall interaction energies are lower (more negative), as expected from the importance of hydrogen bonding and electrostatic interactions of polar residues. It is especially interesting that the P–NP distribution resembles the P–P one more than the NP–NP one.

In pairs with different charge states, those containing at least one positively charged residue and those containing at least one negatively charged residue exhibit comparable energy levels. In contrast, mixed amino-acid pairs containing both positively and negatively charged residues show the lowest interaction energies, with binding strengths significantly stronger than those of all other types, as expected from the strong coulomb attraction.

Overall, polar interactions and charge complementarity are the main driving forces for the formation of strongly bound amino-acid pairs.

3.3. Other test sets

3.3.1. SCAI dataset. The SCAI dataset,[42] developed by Berka *et al.*, is a dataset used to study amino-acid side-chain pair interactions. It covers all 20 types of amino-acid side chain pairs (a total of 400 pairs), from which 24 representative pairs were selected for energy calculations, providing a reliable basis for studying protein side chain interactions. In this study, 23 functionals were used to compute the interaction energies of the 24 amino-acid pair systems in the SCAI dataset. The detailed results are shown in SI Table S9, and Fig. 5A displays the top 15 ranked functionals in this dataset.

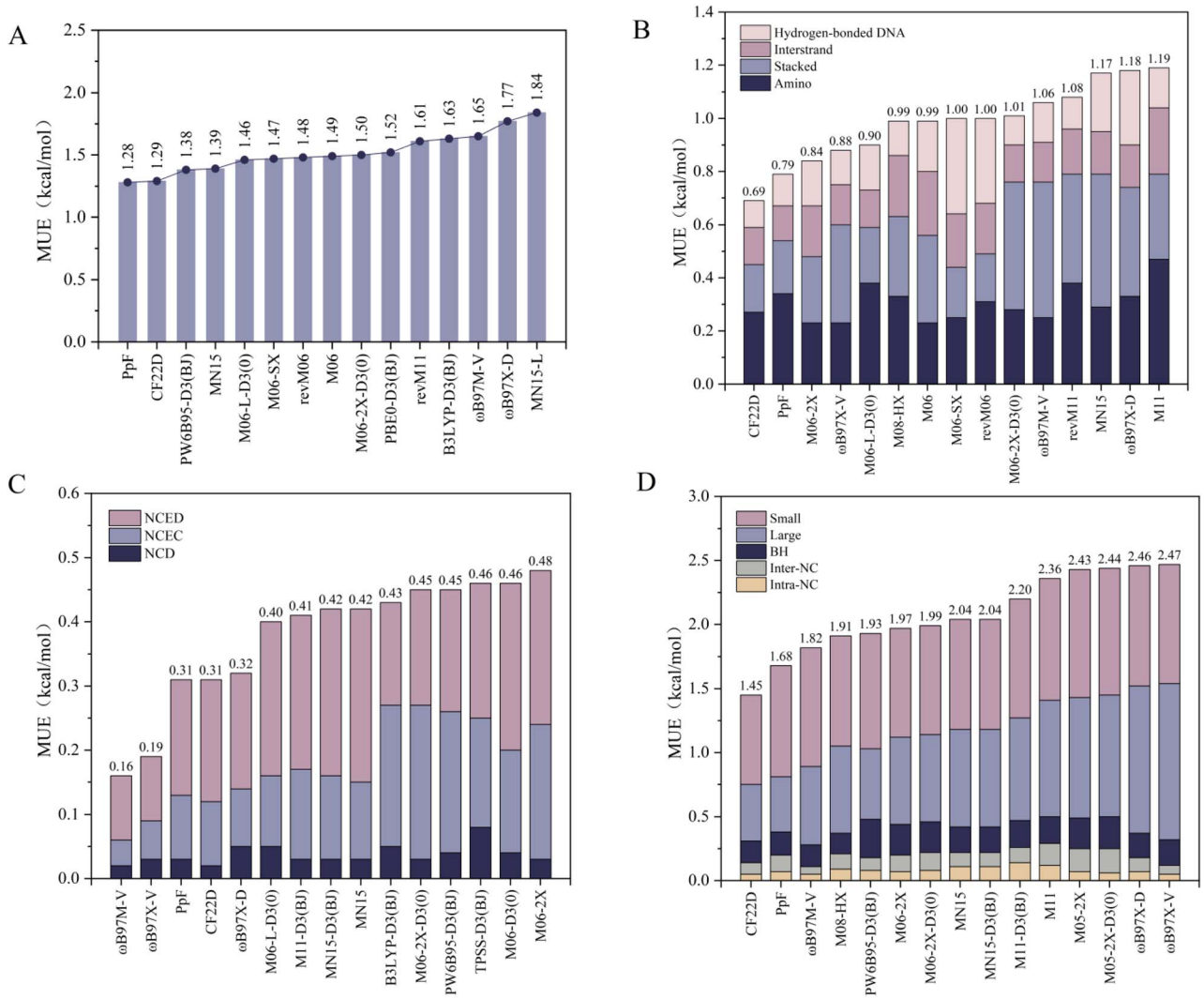

Fig. 5 Results for databases from the literature. (A) MUES (kcal mol^{-1}) of the top 15 functionals on the SCAI dataset. (B) MUES (kcal mol^{-1}) of the top 15 functionals on the JSCH dataset. The numbers above each bar represent the total MUE value for each functional. The different colors represent the contributions to the MUE from each database, including "Hydrogen-bonded DNA" (light pink), "Interstrand" (pink), "Stacked" (blue), and "Amino" (dark blue). (C) MUEs (kcalimol) of the top 15 ranked functionals on the MGCDB84_NC database. The numbers above each bar represent the total MUE value for each functional. The different colors represent the contributions to the MUE from each database, including NCED (pink), NCEC (blue), and NCD (dark blue). (D) MUEs (kcal mol^{-1}) of the top 15 ranked functionals on the GMTKNSS database. The numbers above each bar are the total MUE for each functional. The colors represent the contributions to the MUE from each database, including small (pink), interstrand (pink), large (blue), BH (dark blue), inter-NC (gray), and intra-NC (yellow).

In the SCAI dataset, PpF gives the lowest MUE (1.28 kcal mol^{-1}), ranking first, followed by CF22D (1.29 kcal mol^{-1}), demonstrating their outstanding accuracy in describing amino-acid pair interactions. Both functionals incorporate high-order electron correlation and damped dispersion terms, as well as combining high-order perturbation theory with empirical parameter optimization, significantly improving the accuracy of noncovalent interaction calculations. PW6B95-D3(BJ) and MN15 have MUEs of 1.38 kcal mol^{-1} and 1.39 kcal mol^{-1}, respectively, maintaining relatively low error margins. The M06-L functional, without damped dispersion terms, has a larger error (2.52 kcal mol^{-1}), but after introducing the D3 damped dispersion terms and zero-damping function, the error drops significantly to 1.46 kcal mol^{-1}, showing a marked improvement in performance. ωB97M-V and ωB97X-V, which performed excellently in our amino-acid pair training and test sets, did not perform as well in the SCAI dataset, ranking thirteenth and twentieth, respectively.

3.3.2. JSCH dataset. The JSCH dataset,[43] developed by Jurecka *et al.*, focuses on noncovalent interactions in biomolecules and includes nucleic acid bases, amino acids, and their derivatives. This dataset contains four datasets, three of which correspond to different types of interactions in nucleic acid base pairs: hydrogen-bonded base pairs (38 pairs), interstrand base pairs (32 pairs), and stacked base pairs (54 pairs). The fourth dataset contains amino-acid complexes (19 pairs), and we refer to the four datasets as hydrogen-bonded DNA, Interstrand, Stacked, and Amino, respectively.

In the present work, 24 functionals were used to calculate the interaction energies of the JSCH dataset, and the detailed results are presented in SI Table S10. Fig. 5B displays the 15 best performing functionals for this dataset.

Overall, the PpF functional gave the second best result (0.79 kcal mol^{-1}), just behind CF22D, which had an MUE of 0.69 kcal mol^{-1}. M06-2X (0.84 kcal mol^{-1}) and ωB97X-V (0.88 kcal mol^{-1}) follow closely. The good performance of PpF for this dataset that is much broader than AAIE12 demonstrates that the fine tuning on pairs of amino acids has been accomplished while retaining much of the broad accuracy of CF22D and is a key result in showing the success of fine tuning a foundation model.

Among the four datasets in JSCH, PpF ranked first in the Interstrand dataset (0.58 kcal mol^{-1}), ranked fourth in the Hydrogen-bonded DNA dataset (0.45 kcal mol^{-1}), and did not make the top seven in the remaining two datasets (stacked and amino), with MUEs of 0.53 kcal mol^{-1} and 2.56 kcal mol^{-1}, respectively. However, these values are still much lower than the average MUEs for these datasets (1.25 kcal mol^{-1} and 3.95 kcal mol^{-1}). In the amino dataset, the stabilization energies of amino-acid pairs vary widely, with some complexes having stabilization energies exceeding 100 kcal mol^{-1}. Furthermore, the errors for all functionals are generally larger in the amino dataset, with an average MUE of 3.95 kcal mol^{-1}, while the errors are smallest in the Interstrand dataset, with an average MUE of only 0.86 kcal mol^{-1}.

3.3.3. MGCDB84_NC database. The MGCDB84 database,[44] developed by Mardirossian and Head–Gordon, is a comprehensive database focused on main-group chemistry. Noncovalent interactions and thermochemical property data account for 42% and 24% of the database, respectively. Given that the density functional method PpF developed in this study is fine-tuned for noncovalent interactions, we computed only three noncovalent interaction databases from the MGCDB84 database: noncovalent easy dimers (NCED), noncovalent easy clusters (NCEC), and noncovalent difficult dimers (NCD), totaling 35 datasets and 2708 data points. We refer to the selected data as the MGCDB84_NC database. We used this data to evaluate the performance of various functionals in computing noncovalent interactions. The detailed results are provided in SI Table S11. Fig. 5C displays the results for the 15 best performing functionals in this round of tests.

In terms of overall performance, the MUE for PpF is 0.31 kcal mol^{-1}, ranking third, just behind ωB97M-V (0.16 kcal mol^{-1}) and ωB97X-V (0.19 kcal mol^{-1}). CF22D, MN15-D3(BJ), and MN15 rank fourth, eighth, and ninth, respectively. The improvement of PpF over CF22D, MN15, and MN15-D3(BJ), which have the same functional form, serves as a measure of the progress made through successive supervised learning.

PpF performs relatively well in the NCED and NCEC databases, ranking fourth in both, with MUEs of 0.21 kcal mol^{-1} and 0.84 kcal mol^{-1}, respectively, while CF22D ranks one place behind PpF in both. CF22D performs better in the NCD database, ranking first with an MUE of 0.50 kcal mol^{-1}, while PpF has an MUE of 0.66 kcal mol^{-1}, ranking seventh. This demonstrates the complementarity of the two functionals in addressing noncovalent interaction scenarios.

3.3.4. GMTKN55 database. The GMTKN55 database,[51] integrated by Goerigk *et al.*, covers thermochemical, kinetic, and noncovalent interaction energies of main-group elements. This database contains 1505 data points, divided into five subgroups: small, large, BH, inter-NC and intra-NC. These correspond respectively to basic properties and reaction energies of small molecular systems, reaction energies and isomerization reactions of large molecular systems, reaction barriers, intermolecular noncovalent interactions, and intramolecular noncovalent interactions. In the present study, 38 mainstream density functionals were compared for their performance on the GMTKN55 database, with detailed results presented in SI Table S12.

The first six functionals in SI Table S12 are five doubly hybrid[52] functionals and the machine-learned DM21 (ref. 53) functional; due to the significantly higher computational cost of these six functionals, we use only the other 30 functionals, in the rankings discussed in this section, and we refer to these 30 functionals as the ranked functionals. Fig. 5D displays the top 15 ranked functionals in this database.

Overall, CF22D outperforms all ranked functionals on the entire database, with an MUE of 1.45 kcal mol^{-1}, ranking first. Following closely is PpF, with an MUE of 1.68 kcal mol^{-1}, demonstrating high computational accuracy on this broad dataset and again demonstrating that fine tuning on pairs of amino acids has been accomplished while retaining much of the broad accuracy of CF22D. The MUE of PpF is 1.01 kcal mol^{-1} lower than the overall average MUE. The ωB97M-V and M08-HX functionals ranked third and fourth among the ranked functionals, demonstrating high accuracy without damped dispersion terms.

At the request of a reviewer, we examined the performance of three versions of the SCAN functional. The regularized functionals r2SCAN (3.13 kcal mol^{-1}) and r2SCAN-3c (2.81 kcal mol^{-1}) show improved performance compared with SCAN-D3(0) (3.35 kcal mol^{-1}), but the overall performance of all three versions is inferior to that of most of the other functionals.

Among the datasets, PpF ranked first in the large dataset (2.67 kcal mol^{-1}), fourth in the BH dataset, and fifth in the small dataset. However, the performance of PpF in the inter-NC and intra-NC datasets was somewhat lower, with MUEs of 0.64 kcal mol^{-1} and 0.37 kcal mol^{-1}, respectively. Although its rankings were lower, the MUEs are still much lower than the average MUEs for these datasets (0.87 kcal mol^{-1} and 0.53 kcal mol^{-1}). This suggests that although the PpF functional performs moderately for some more complex noncovalent interactions, it still demonstrates strong computational accuracy on most datasets, particularly in the Large dataset, where its performance is comparable to some doubly hybrid functionals.

The results of the doubly hybrid functionals in SI Table S12 illustrates that adding orbital-based nonlocal correlation terms can improve the results. More generally, adding additional ingredients is an effective way to enhance functional performance. However, among the five doubly hybrid functionals,

only one has an MUE lower than PpF by more than 0.39 kcal mol^{-1}, while two have higher MUEs. SI Table S12 also indicates that the deep-learning functional DM21 performs better, but only slightly better, than PpF.

4. Conclusions

The PpF density functional is specifically designed to accurately compute noncovalent interaction energies between amino-acid pairs such as those in protein–peptide complexes. Developed through a performance-triggered iterative supervised training strategy that optimizes the functional in the presence of a damped dispersion term, PpF achieves excellent accuracy and stability for amino-acid pair interactions. PpF outperforms existing methods on the AAIE12, AAIE24 and external SCAI dataset, and ranks highly in the JSCH, GMTKN55, and MGCDB84_NC benchmark databases, demonstrating its broad applicability to biomolecular and general noncovalent problems.

Along with the recent significant progress in functional development and dataset construction, there remain many avenues for future research. Further leveraging the advantages of data-driven approaches to density functional theory is very promising.[54–57] For example, building on the present work, further optimization of the PpF functional is possible by broadening the target data to include more complex noncovalent interactions, such as those controlling conformational dynamics and many-body effects in the protein–peptide binding process. Another promising step forward would be to expand the dataset by incorporating protein–peptide interaction pairs related to major diseases, in order to support rational drug design and disease mechanism research.[58]

The AAPS260 and AAIE260 datasets, constructed based on the PpF functional developed in this study, provide extensive structural and energy information for amino-acid interaction research and the validation of computational methods. The AAIE260 dataset can serve as a benchmark for evaluating the performance of various computational methods in predicting amino-acid interaction energies.

The PpF functional developed in this study has good prospects for broad applications. It can be used to optimize the geometries of systems containing amino acids, to calculate the interaction energies of amino-acid pairs, to explore mechanism, and to perform direct dynamics simulations. As compared to traditional functionals, the PpF functional shows better adaptability to complex biomolecular environments and is expected to provide strong theoretical support for binding affinity prediction, drug design, and studies of protein function regulation and molecular recognition.

Author contributions

Yini Zhou: data curation, formal analysis, investigation, methodology, software, validation, visualization, writing – original draft. Tao Li: software, validation, funding acquisition. Yaqi Li: data curation. Jianda Yue: methodology. Qifeng Tian: validation. Zhonghua Liu: resources. Donald G. Truhlar: funding acquisition, supervision, writing – review & editing. Ying Wang: conceptualization, funding acquisition, project administration, resources, supervision, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Data availability

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d6sc01551d.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Grants No. 22473041 to Y.W.), the Natural Science Foundation of Hunan Province (Grant No. 2024JJ2042 to Y. W), the Scientific research project of Education Department of Hunan Province (Key Project, Grant No. 23A0084 to Y. W.), and the 2024 Hunan Provincial Graduate Research and Innovation (Grant No. CX20240544 to T. L.), and the Air Force Office of Scientific Research (grant FA9550-20-1-0360 to D. T.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank the Bioinformatics Center of Hunan Normal University for providing computer resources.

References

1 M. Muttenthaler, G. F. King, D. J. Adams and P. F. Alewood, Trends in peptide drug discovery, *Nat. Rev. Drug Discovery*, 2021, **20**, 309–325.
2 Y. Lei, S. Li, Z. Liu, F. Wan, T. Tian, S. Li, D. Zhao and J. Zeng, A deep-learning framework for multi-level peptide–protein interaction prediction, *Nat. Commun.*, 2021, **12**, 5465.
3 T. Tsaban, J. K. Varga, O. Avraham, Z. Ben-Aharon, A. Khramushin and O. Schueler-Furman, Harnessing protein folding neural networks for peptide–protein docking, *Nat. Commun.*, 2022, **13**, 176.
4 P. Campitelli, T. Modi, S. Kumar and S. B. Ozkan, The role of conformational dynamics and allostery in modulating protein evolution, *Annu. Rev. Biophys.*, 2020, **49**, 267–288.
5 S. M. P. Vadevoo, S. Gurung, H.-S. Lee, G. R. Gunassekaran, S.-M. Lee, J.-W. Yoon, Y.-K. Lee and B. Lee, Peptides as multifunctional players in cancer therapy, *Exp. Mol. Med.*, 2023, **55**, 1099–1109.
6 A. Karshikoff, *Non-covalent interactions in proteins*, World Scientific, 2006.
7 A. Baruch Leshem, S. Sloan-Dennison, T. Massarano, S. Ben-David, D. Graham, K. Faulds, H. E. Gottlieb, J. H. Chill and A. Lampel, Biomolecular condensates formed by designer minimalistic peptides, *Nat. Commun.*, 2023, **14**, 421.
8 B. Jawad, P. Adhikari, R. Podgornik and W.-Y. Ching, Key interacting residues between RBD of SARS-CoV-2 and ACE2 receptor: combination of molecular dynamics simulation

and density functional calculation, *J. Chem. Inf. Model.*, 2021, **61**, 4425–4441.

9 W. Dawson, A. Degomme, M. Stella, T. Nakajima, L. E. Ratcliff and L. Genovese, Density functional theory calculations of large systems: Interplay between fragments, observables, and computational complexity, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1574.

10 K. Tsuboyama, J. Dauparas, J. Chen, E. Laine, Y. Mohseni Behbahani, J. J. Weinstein, N. M. Mangan, S. Ovchinnikov and G. J. Rocklin, Mega-scale experimental analysis of protein folding stability in biology and design, *Nature*, 2023, **620**, 434–444.

11 P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.*, 1964, **136**, B864.

12 W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, 1965, **140**, A1133.

13 G. E. Scuseria and V. N. Staroverov, in *Theory and applications of computational chemistry*, Elsevier, 2005, pp. 669–724.

14 H. S. Yu, S. L. Li and D. G. Truhlar, Perspective: Kohn-Sham density functional theory descending a staircase, *J. Chem. Phys.*, 2016, 145.

15 M. Bogojeski, L. Vogt-Maranto, M. E. Tuckerman, K.-R. Müller and K. Burke, Quantum chemical accuracy from density functional approximations via machine learning, *Nat. Commun.*, 2020, **11**, 5223.

16 J.-D. Chai and M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

17 N. F. Bras, M. A. Perez, P. A. Fernandes, P. J. Silva and M. J. Ramos, Accuracy of density functionals in the prediction of electronic proton affinities of amino acid side chains, *J. Chem. Theory Comput.*, 2011, **7**, 3898–3908.

18 Y. Zhao and D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theor. Chem. Acc.*, 2008, **120**, 215–241.

19 P. Wang, C. Shu, H. Ye and M. Biczysko, Structural and energetic properties of amino acids and peptides benchmarked by accurate theoretical and experimental data, *J. Phys. Chem.*, 2021, **125**, 9826–9837.

20 S. Damodaran and K. L. Parkin, in *Fennema's food chemistry*, CRC press, 2017, pp. 235–356.

21 Y. Liu, C. Zhang, Z. Liu, D. G. Truhlar, Y. Wang and X. He, Supervised learning of a chemistry functional with damped dispersion, *Nat. Comput. Sci.*, 2023, **3**, 48–58.

22 Y.-Q. Zhang, Y. Xia and B. Long, Quantitative kinetics for the atmospheric reactions of Criegee intermediates with acetonitrile, *Phys. Chem. Chem. Phys.*, 2022, **24**, 24759–24766.

23 Y. Li, Y. Wang, R. M. Zhang, X. He and X. Xu, Comprehensive theoretical study on four typical intramolecular hydrogen shift reactions of peroxy radicals: multireference character, recommended model chemistry, and kinetics, *J. Chem. Theory Comput.*, 2023, **19**, 3284–3302.

24 A. Gonzalez-Lafont, T. N. Truong and D. G. Truhlar, Direct dynamics calculations with NDDO (neglect of diatomic differential overlap) molecular orbital theory with specific reaction parameters, *J. Phys. Chem.*, 1991, **95**, 4618–4627.

25 I. Rossi and D. G. Truhlar, Parameterization of NDDO wavefunctions using genetic algorithms. An evolutionary approach to parameterizing potential energy surfaces and direct dynamics calculations for organic reactions, *Chem. Phys. Lett.*, 1995, **233**, 231–236.

26 J. Pu and D. G. Truhlar, Parametrized direct dynamics study of rate constants of H with CH 4 from 250 to 2400 K, *J. Chem. Phys.*, 2002, **116**, 1468–1478.

27 A. S. Hauser and B. r. Windshügel, LEADS-PEP: a benchmark data set for assessment of peptide docking performance, *J. Chem. Inf. Model.*, 2016, **56**, 188–200.

28 K. B. Santos, I. A. Guedes, A. L. Karl and L. E. Dardenne, Highly flexible ligand docking: Benchmarking of the DockThor program on the LEADS-PEP protein–peptide data set, *J. Chem. Inf. Model.*, 2020, **60**, 667–683.

29 H. Cheng, G.-G. Wang, L. Chen and R. Wang, A dual-population multi-objective evolutionary algorithm driven by generative adversarial networks for benchmarking and protein-peptide docking, *Comput. Biol. Med.*, 2024, **168**, 107727.

30 Y. Masoudi-Sobhanzadeh, B. Jafari, S. Parvizpour, M. M. Pourseif and Y. Omidi, A novel multi-objective metaheuristic algorithm for protein-peptide docking and benchmarking on the LEADS-PEP dataset, *Comput. Biol. Med.*, 2021, **138**, 104896.

31 D. W. Zhang and J. Zhang, Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy, *J. Chem. Phys.*, 2003, **119**, 3599–3605.

32 C. Tian, K. Kasavajhala, K. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay and Q. Wu, ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution, *J. Chem. Theory Comput.*, 2019, **16**, 528–552.

33 E. G. Hohenstein and C. D. Sherrill, Wavefunction methods for noncovalent interactions, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 304–326.

34 J. Rezac and P. Hobza, Benchmark calculations of interaction energies in noncovalent complexes and their applications, *Chem. Rev.*, 2016, **116**, 5038–5071.

35 K. Patkowski, Benchmark databases of intermolecular interaction energies: Design, construction, and significance, *Annu. Rep. Comput. Chem.*, 2017, **13**, 3–91.

36 C. Riplinger and F. Neese, An efficient and near linear scaling pair natural orbital based local coupled cluster method, *J. Chem. Phys.*, 2013, 138.

37 A. Hellweg and D. Rappoport, Development of new auxiliary basis functions of the Karlsruhe segmented contracted basis sets including diffuse basis functions (def2-SVPD, def2-TZVPPD, and def2-QVPPD) for RI-MP2 and RI-CC calculations, *Phys. Chem. Chem. Phys.*, 2015, **17**, 1010–1017.

38 J. Zheng, X. Xu and D. G. Truhlar, Minimally augmented Karlsruhe basis sets, *Theor. Chem. Acc.*, 2011, **128**, 295–305.
39 S. Grimme, S. Ehrlich and L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
40 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, *J. Chem. Phys.*, 1980, **72**, 650–654.
41 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**, 154104.
42 K. Berka, R. Laskowski, K. E. Riley, P. Hobza and J. Vondrasek, Representative amino acid side chain interactions in proteins. A comparison of highly accurate correlated ab initio quantum chemical and empirical potential procedures, *J. Chem. Theory Comput.*, 2009, **5**, 982–992.
43 P. Jurečka, J. Šponer, J. Černý and P. Hobza, Benchmark database of accurate (MP2 and CCSD (T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.
44 N. Mardirossian and M. Head-Gordon, Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals, *Mol. Phys.*, 2017, **115**, 2315–2372.
45 P. Verma and D. G. Truhlar, *Geometries for Minnesota database 2019*, 2019.
46 F. Neese, Software update: The ORCA program system—Version 5.0, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1606.
47 M. Frisch, *gaussian 09, Revision d. 01*, Gaussian, Inc, *Wallingford CT*, 2009, 201.
48 Y. Zhao, R. Peverati, K. Yang and D. G. Truhlar, *MN-GFM, version 6.3: Minnesota Gaussian Functional Module*, University of Minnesota, Minneapolis, 2012, 45.
49 M. e. Frisch, G. Trucks, H. B. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, G. Petersson and H. Nakatsuji, *Gaussian 16*, 2016.
50 D. G. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio and A. Alenaizan, PSI4 1.4: Open-source software for high-throughput quantum chemistry, *J. Chem. Phys.*, 2020, **152**, 184108.
51 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.
52 Y. Zhao, B. J. Lynch and D. G. Truhlar, Doubly hybrid meta DFT: New multi-coefficient correlation and density functional methods for thermochemistry and thermochemical kinetics, *J. Phys. Chem.*, 2004, **108**, 4786–4791.
53 J. Kirkpatrick, B. McMorrow, D. H. Turban, A. L. Gaunt, J. S. Spencer, A. G. Matthews, A. Obika, L. Thiry, M. Fortunato and D. Pfau, Pushing the frontiers of density functionals by solving the fractional electron problem, *Science*, 2021, **374**, 1385–1389.
54 S. Shanker and M. F. Sanner, Predicting protein–peptide interactions: Benchmarking deep learning techniques and a comparison with focused docking, *J. Chem. Inf. Model.*, 2023, **63**, 3158–3170.
55 X. Xu and X. Zou, Predicting protein–peptide complex structures by accounting for peptide flexibility and the physicochemical environment, *J. Chem. Inf. Model.*, 2021, **62**, 27–39.
56 S. Romero-Molina, Y. B. Ruiz-Blanco, J. Mieres-Perez, M. Harms, J. Munch, M. Ehrmann and E. Sanchez-Garcia, PPI-affinity: A web tool for the prediction and optimization of protein–peptide and protein–protein binding affinity, *J. Proteome Res.*, 2022, **21**, 1829–1841.
57 S. A. Spronk, Z. L. Glick, D. P. Metcalf, C. D. Sherrill and D. L. Cheney, A quantum chemical interaction energy dataset for accurately modeling protein-ligand interactions, *Sci. Data*, 2023, **10**, 619.
58 H. Wang, R. S. Dawber, P. Zhang, M. Walko, A. J. Wilson and X. Wang, Peptide-based inhibitors of protein–protein interactions: biophysical, structural and cellular consequences of introducing a constraint, *Chem. Sci.*, 2021, **12**, 5977–5993.