

ARTICLE OPEN

A high-throughput framework for determining adsorption energies on solid surfaces

Joseph H. Montoya¹ and Kristin A. Persson^{1,2}

In this work, we present a high-throughput workflow for calculation of adsorption energies on solid surfaces using density functional theory. Using open-source computational tools from the Materials Project infrastructure, we automate the procedure of constructing symmetrically distinct adsorbate configurations for arbitrary slabs. These algorithms are further used to construct and run workflows in a standard, automated way such that user intervention in the simulation procedure is minimal. To validate our approach, we compare results from our workflow to previous experimental and theoretical benchmarks from the CE27 database of chemisorption energies on solid surfaces. These benchmarks also illustrate how the task of performing and managing over 200 individual density functional theory calculations may be reduced to a single submission procedure and subsequent analysis. By enabling more efficient high-throughput computations of adsorption energies, these tools will accelerate theory-guided discovery of advanced materials for applications in catalysis and surface science.

npj Computational Materials (2017)3:14; doi:10.1038/s41524-017-0017-z

INTRODUCTION

With applications in heterogeneous catalysis, medical imaging, and microelectronics, surface science has been and will likely continue to be key to developing advanced technologies. In recent years, electronic structure simulations using density functional theory (DFT) have greatly enhanced insights into the chemical properties of solid surfaces.¹ These calculations have even been used to make successful predictions of new materials with desirable catalytic properties for such processes as methanol synthesis² and electrochemical hydrogen evolution.³

Despite these successes, most modern DFT-based surface science uses an approach that relies heavily on human intuition to perform and tune individual calculations of individual surfaces and adsorbates. This reliance on human intuition stems partially from the complexity of the electronic structure calculations themselves, since ideal parameters specific to slabs and surfaces may differ from those used in bulk structure and property determination. However, recent efforts to perform bulk calculations in a high-throughput manner have demonstrated that many of the intuitive aspects of determining solid properties like thermodynamic stability,⁴ elastic properties,⁵ and surface energy⁶ using DFT can be effectively automated.

Nevertheless, the intuition required for successfully performing surface slab and adsorption energy calculations is not purely that of tuning electronic structure parameters. Providing initial guesses for adsorbate structures based on bonding geometries and selecting the sites for consideration of adsorption are also key steps in most manual surface science workflows, and can be difficult to automate in a comprehensive way. These tasks essentially amount to a human pre-processing of slab geometries, and are integral to ensuring that the adsorption energies that best represent a given surface facet's chemical reactivity are selected for further modeling of such properties as catalytic activity or

theoretical overpotential. Global optimization of adsorbate structures using constrained minima hopping,⁷ metadynamics,⁸ or Bayesian optimization⁹ may be used to exhaustively sample the potential energy surface of a given system, but the computational cost of these methods makes them unwieldy for a high-throughput approach.

In this work, we present a workflow for performing DFT calculations of slabs and adsorbed species by which high-throughput operation might be achieved. In particular, we present algorithms and tools for the automation of adsorption geometry determination from structural properties of slabs and their associated bulk structures. Furthermore, we present a standard workflow using the automation tools of the Materials Project that may be used to generate adsorption data from DFT in a high-throughput manner. As an initial example, we demonstrate the workflow by comparing its results to a data set from the CE27 chemisorption energy benchmarks.¹⁰ In addition to validating our methods, these benchmarks illustrate how our workflow can be used to flexibly handle a diverse set of adsorbates, slabs, and bulk structures without explicit specification of adsorption geometries or computational parameters for individual jobs, reducing the management of over 200 DFT calculations (and potentially one to two orders of magnitude more) to that of a single submission.

RESULTS AND DISCUSSION

We present benchmarked examples of how our workflow may be used to generate first-principles adsorption energy data in high-throughput. Our benchmarks are based on a test set that corresponds to the CE27 database of chemisorption energies. These data include chemisorption energies of H₂, N₂, CO, O, and NO on various low-index facets of elemental crystals. In our benchmarking, we compare data generated from our automated workflow to experimental chemisorption values from this

¹Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA and ²Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA

Correspondence: Kristin A. Persson (kristinpersson@berkeley.edu)

Received: 14 December 2016 Revised: 21 February 2017 Accepted: 28 February 2017

Published online: 30 March 2017

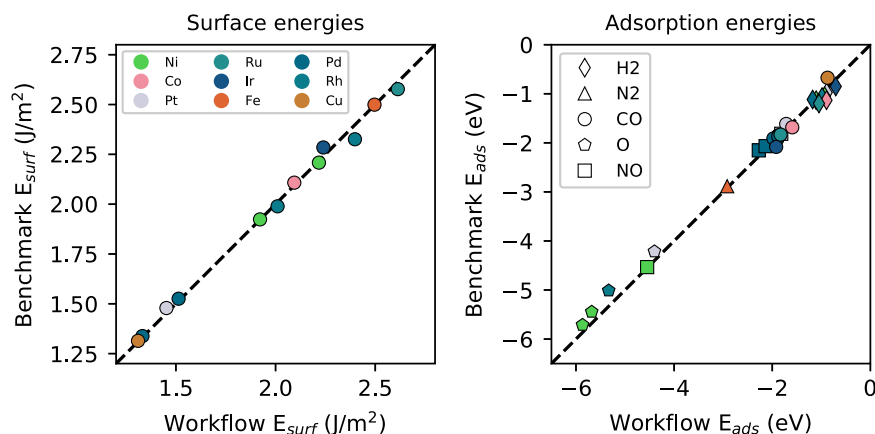


Fig. 1 Computational benchmarks, in which we compare results from our workflow to previously computed adsorption and surface energies. Surface energy benchmarks are from Tran *et al.*⁶ Note that our workflow may not include surface reconstructions accounted for in the benchmarking set, which may account for small deviations. Calculated adsorption energy benchmarks are from Wellendorf *et al.*¹⁰ and include chemisorption energies corresponding to materials and crystal facets featured in the CE27 database. Data for both chemisorption and surface energy benchmarks are from calculations using the PBE functional. A table is provided in the [Supplementary Information](#) that includes details on surface facets and data for surface and adsorption energies

database, as well as previously computed values from two functionals commonly used to calculate adsorption energies.

Since clean slab calculations are necessary reference states in adsorption energy calculations, data from our workflow may also be used to estimate surface energies, which we benchmark with recent work cataloging the surface energies of all elemental crystals.⁶ The resultant data closely matches the benchmarks from previous calculations, as shown in Fig. 1a. These calculations are essentially identical in methodology to our workflow, and thus are very closely replicated with a mean-absolute error of 0.02 eV. Small deviations in this benchmarking comparison may be due to surface reconstructions, which are treated more thoroughly in the benchmarking study.

To benchmark adsorption energies, we refer to previous work from Wellendorf *et al.* intended to catalog a variety of density functionals for surface science.¹⁰ These results are also shown in Fig. 1b, and compare to calculation benchmarks with a mean-absolute-error of 0.2 eV. Due to limitations on the reported details of the procedure used benchmarking study, it is not precisely clear which elements of our workflow differ from that used in the computational benchmarking set. However, we conjecture that the observed discrepancies primarily arise from differences in optimization routines, DFT formalisms, and the different pseudopotentials of the respective codes (GPAW^{11, 12} vs. Vienna ab-initio software package (VASP)). In addition, treatment of reference states, which in our study are the uncorrected electronic energies of the precursor molecule, may account for small systematic deviations most notably observed in the oxygen adsorption energies. Ultimately, however, the comparison seems robust enough to capture trends in chemisorption behavior within DFT error and therefore further establishes our workflow's reliability.

We further catalog our initial results as benchmarked against the CE27 database of experimental chemisorption energies.^{13–21} We note here that changing the input parameters of all DFT calculations is very simply achieved by the workflow modification tools present in atomate, which we have used to run two identical workflows with the Perdew-Burke-Ernzerhof (PBE)²² and revised Perdew-Burke-Ernzerhof (RPBE)²³ functionals. As previously reported,^{10, 13, 23} our benchmarks in Fig. 2 show that results using RPBE are in much closer quantitative agreement with experiment and that PBE consistently underpredicts the chemisorption energies of molecules. Both functionals consistently reproduce trends in the calculated adsorption energies compared with experiment.

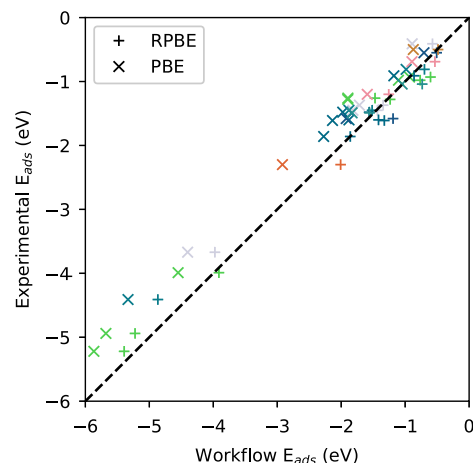


Fig. 2 Experimental benchmarks, in which we compare results from our workflow using both the PBE and RPBE functionals to experimental chemisorption energies from the CE27 database

These examples evidence how our high-throughput approach may be adapted for the determination of adsorption energies. However, our infrastructure in its current form still has limitations, which are the foci of future improvements. Complex adsorbate molecules, such as H_2O , HOO^* , or C_6H_6 , may require manual intervention in workflow generation to include molecular configurations accounting for rotational degrees of freedom. Distinct pairings of rotational configurations of molecules with surfaces might be generated using similar geometric analysis of molecular symmetry, but will likely remain complex. In addition, symmetrically distinct adsorption sites on stepped and kinked surfaces, particularly important to catalytic activity, are more numerous than those that are typically accounted for in many human workflows. Lastly, we note that kinetic barriers along reaction pathways corresponding to bond breaking or formation and coverage effects are key parameters in catalysis, and are not yet accounted for in our approach. These, along with efforts toward further understanding structure–property relationships relevant to adsorption energies, are the subject of ongoing efforts. Recent machine-learning approaches²⁴ and more advanced structural

descriptors²⁵ have been promising for improving understanding of adsorption behavior. These approaches may also help to understand which configurations and adsorption sites are most likely representative of the most stable or catalytically active, and thus they may help streamline the process of simulating these phenomena.

METHODS

In this work, we introduce an algorithm for finding the adsorption sites on an arbitrary surface. The algorithm initializes with a selection of the “surface sites”, which can be designated manually, selected using a threshold distance window from the largest extent of the slab along the miller index, or selected by determining whether a given site is undercoordinated relative to its bulk counterpart according to a Voronoi coordination determination implemented in the pymatgen open-source software.²⁶ From this set of surface sites, a 2D Voronoi tessellation of their coordinates (plus those in the adjacent periodic images from the slab) projected onto the plane perpendicular to the miller index is calculated. “On-top” sites are assigned to the surface sites themselves, while “bridge” sites are assigned to the midpoints of the edges of the Voronoi tessellation. As a result, “hollow” sites are assigned to the center of the ensemble of sites that comprise a Voronoi face in the tessellation. Any sites generated outside of the unit cell from the extended surface mesh are translated such that they are placed inside the slab unit cell.

The resultant sites are then filtered from two criteria. Sites within a certain distance of another site, based on user input with a default of 0.1 Å, are discarded. Default operation of the algorithm then identifies sites among these which are symmetrically equivalent according to the symmetry operations of the slab structure. Ultimately, these filters yield a set of symmetrically and geometrically distinct sites for an arbitrary slab, contingent on the appropriate selection of surface sites. In Fig. 3a–d, we show the result of progressive steps of our algorithm with a simple example on Ni 111, for which the on-top, bridge, fcc (hollow), and hollow sites are properly generated. We provide further examples of this algorithm’s output and operation for alternate structures (including binary and ternary oxides) and surfaces in the [Supplementary Information](#).

From this set of adsorption sites, we generate a workflow similar to those previously constructed for structure optimization and elastic tensor determination.⁵ This workflow is generated using the atomate code package, which combines the open-source FireWorks²⁷ and pymatgen²⁶ codes to create standard workflows for computational analysis of materials. More specifically, the workflow we present herein begins with a standard structure optimization in the VASP^{28, 29} to ensure the lattice constant is converged within the user’s DFT parameters. This also allows the structure to be optimized using a different parameter set than what is prescribed in the standard materials project workflow, if desired. Previous workflows designed with a similar purpose, i.e., conducting calculations that derive properties from an initial calculation, have employed dynamic workflows in the FireWorks package to generate new tasks on the fly during workflow operation. However, in our experience, this approach presents maintenance difficulties, since the dynamic workflows are often difficult to

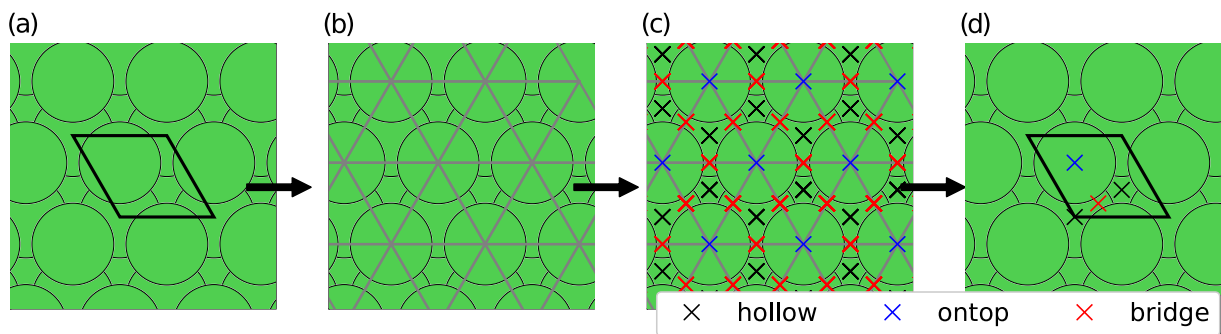


Fig. 3 Adsorption site selection for the Ni (111) slab, in which a periodic slab shown in (a) has its surface sites selected to generate a Delaunay triangulation network (b), with which adsorption sites are placed at face edges, centers, and vertices (c) and filtered such that symmetrically equivalent and very close sites are removed to yield a minimal set of distinct adsorption sites (d)

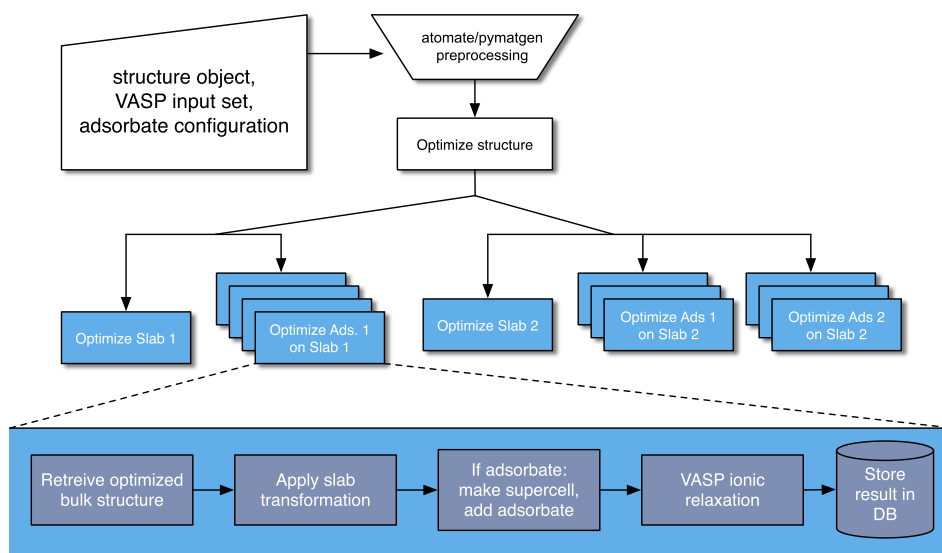


Fig. 4 Atomate workflow for calculating adsorption energies: this workflow is generated in the atomate software package from structure, VASP parameters, and adsorbate configuration inputs. The workflow begins with a stress-based structure optimization that branches into ionic relaxation of slab and adsorbate geometries. In each task, the results of the VASP simulation are either stored in a database or output to the filesystem in a JSON document

debug if problems arise in the spawned tasks or in the initial tasks that they are derived from. As such, we opt for a construction of a fixed number of tasks at the outset of workflow construction. This is achieved by analyzing the unoptimized input structure using the previously described algorithms and storing each of the transformations necessary to construct slabs and adsorbate structures as functions that take the optimized structure resulting from the initial task as input.

Each FireWork implemented in the workflow conducts a preprocessing step by which any system-specific parameters (e.g., parallelization or algorithmic settings) are applied to the VASP input files, a VASP optimization step run via the custodian job management framework that can correct standard errors on the fly, and then a post-processing step by which the results of the DFT simulation are collected, stored in a JavaScript Object Notation (JSON) document, which may be uploaded to an external database or output to the local filesystem. In Fig. 4, we sketch the general structure of our workflow. Inputs to the surface absorption workflow include a bulk structure, VASP parameters, and an “adsorbate configuration”, which supplies the chemical identity and geometry of the adsorbate in addition to the miller indices that adsorbate is to be placed on. Ultimately, this allows a user to generate an entire workflow from a minimal set of input parameters.

In the supplement, we include code used to generate both the workflow used to calculate the benchmarking data from the CE27. In this example, a specific facets of various metal surfaces corresponding specifically to those included in the CE27 database are included by using the appropriate atomate and pymatgen input parameters. However, we also note that the workflow outlined herein can be used in a much more flexible way to explore every distinct facet of a large variety of materials subject to user-supplied constraints. To illustrate this, we include a further example in the supplement that generates a workflow in which the binding energies for oxygen evolution intermediates on various complex materials are calculated for each distinct low-index facet of a given material, which include distinct terminations of a given surface. The capacity of our approach to achieve this more clearly distinguishes it from traditional manual workflows in computational surfaces sciences that typically confine themselves to easily constructed slabs or structural motifs, which may be neither the most stable nor the most catalytically active.

In the [Supplementary Information](#), we provide IPython notebooks with the requisite package installation instructions and examples that may be used to generate the structures and workflows involved in the two benchmarking examples.

ACKNOWLEDGEMENTS

This work was intellectually led by the Materials Project Center, supported by the DOE Basic Energy Sciences Grant No. EDCBEE. Benchmarking calculations presented in this work used resources of the National Energy Research Scientific Computing Center (NERSC), supported by the Office of Basic Energy Sciences of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors also would like to acknowledge prior work from Kiran Mathew, Anubhav Jain, Shyam Dwaraknath, and Muratahan Aykol on the atomate package used to implement the workflow outlined herein, which will be featured in an upcoming publication.

AUTHOR CONTRIBUTIONS

J.M. performed the workflow design, code implementation, and calculation analysis. K.P. is the primary investigator and supervised the workflow and code development. Both authors contributed to the writing and editing of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

- Nørskov, J. K., Bligaard, T., Rossmeisl, J. & Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* **1**, 37–46 (2009).
- Studt, F. *et al.* Discovery of a Ni-Ga catalyst for carbon dioxide reduction to methanol. *Nat. Chem.* **6**, 320–324 (2014).
- Hinnemann, B. *et al.* Biomimetic hydrogen evolution: MoS₂ nanoparticles as catalyst for hydrogen evolution. *J. Am. Chem. Soc.* **127**, 5308–5309 (2005).
- Jain, A. *et al.* Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 11002 (2013).
- de Jong, M. *et al.* Charting the complete elastic properties of inorganic crystalline compounds. *Sci. Data* **2**, 150009 (2015).
- Tran, R. *et al.* Surface energies of elemental crystals. *Sci. Data* **3**, 160080 (2016).
- Peterson, A. A. Global optimization of adsorbate–surface structures while preserving molecular identity. *Top. Catal.* **57**, 40–53 (2013).
- Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12562–12566 (2002).
- Carr, S., Garnett, R. & Lo, C. BASC: applying Bayesian optimization to the search for global minima on potential energy surfaces. In *Proceedings of the 33rd International Conference on Machine Learning* 898–907 (2016).
- Wellendorff, J. *et al.* Density functionals for surface science: exchange–correlation model development with Bayesian error estimation. *Phys. Rev. B* **85**, 235149 (2012).
- Mortensen, J. J., Hansen, L. B. & Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B* **71**, 035109 (2005).
- Enkovaara, J. *et al.* Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys. Condens. Matter* **22**, 253202 (2010).
- Abild-Pedersen, F. & Andersson, M. P. CO adsorption energies on metals with correction for high coordination adsorption sites—a density functional study. *Surf. Sci.* **601**, 1747–1753 (2007).
- Brown, W. A., Kose, R. & King, D. A. Femtomole adsorption calorimetry on single-crystal surfaces. *Chem. Rev.* **98**, 797–832 (1998).
- Conrad, H., Ertl, G., Koch, J. & Latta, E. E. Adsorption of CO on Pd single crystal surfaces. *Surf. Sci.* **43**, 462–480 (1974).
- Behm, R. J., Christmann, K., Ertl, G. & Van Hove, M. A. Adsorption of CO on Pd(100). *J. Chem. Phys.* **73**, 2984–2995 (1980).
- Ge, Q., Kose, R. & King, D. A. Adsorption energetics and bonding from femtomole calorimetry and from first principles theory. *Adv. Catal.* **45**, 207–259 (2000).
- Fiorin, V., Borthwick, D. & King, D. A. Microcalorimetry of O₂ and NO on flat and stepped platinum surfaces. *Surf. Sci.* **603**, 1360–1364 (2009).
- Yeo, Y. Y., Vattuone, L. & King, D. A. Calorimetric investigation of NO and CO adsorption on Pd{100} and the influence of preadsorbed carbon. *J. Chem. Phys.* **106**, 1990–1996 (1997).
- Bozso, F., Ertl, G., Grunze, M. & Weiss, M. Interaction of nitrogen with iron surfaces: I. Fe(100) and Fe(111). *J. Catal.* **49**, 18–41 (1977).
- Christmann, K. Interaction of hydrogen with solid surfaces. *Surf. Sci. Rep.* **9**, 1–163 (1988).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7420 (1999).
- Ulissi, Z. W., Singh, A. R., Tsai, C. & Nørskov, J. K. Automated discovery and construction of surface phase diagrams using machine learning. *J. Phys. Chem. Lett.* **7**, 3931–3935 (2016).
- Calle-Vallejo, F., Loffreda, D., Koper, M. T. M. & Sautet, P. Introducing structural sensitivity into adsorption–energy scaling relations by means of coordination numbers. *Nat. Chem.* **7**, 403–410 (2015).
- Ong, S. P. *et al.* Python materials genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
- Jain, A. *et al.* FireWorks: a dynamic workflow system designed for high-throughput applications. *Concurr. Comput. Pract. Exp.* **27**, 5037–5059 (2015).
- Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
- Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the *npj Computational Materials* website (doi:10.1038/s41524-017-0017-z).