

[Artificia](https://doi.org/10.1016/j.ailsci.2023.100077)l[IntelligenceintheLifeSciences3(2023)100077](https://doi.org/10.1016/j.ailsci.2023.100077)

|  |  |  |
| --- | --- | --- |
|  | Contentslistsavailableat[ScienceDirect](http://www.ScienceDirect.com) |  |
| ArtificialIntelligenceintheLifeSciences |
| journalhomepage:[www.elsevier.com/locate/ailsci](http://www.elsevier.com/locate/ailsci) |

TrendsandchallengesinchemoinformaticsresearchinLatinAmerica

JazmínMiranda-Salasa,CarlosPeña-Varasb,c,IgnacioValenzuelaMartínezb,   
DionisioA.Olmedod,WilliamJ.Zamorae,f,i,MiguelAngelChávez-Fumagallig,   
DanielaQ.Azevedoh,RachelOliveiraCastilhoh,ViniciusG.Maltarolloh,DavidRamírezb, José L.Medina-Francoa,∗

a*DIFACQUIMResearchGroup,DepartmentofPharmacy,SchoolofChemistry,UniversidadNacionalAutónomadeMéxico,AvenidaUniversidad3000,México04510, Mexico*   
b*DepartamentodeFarmacología,FacultaddeCienciasBiológicas,UniversidaddeConcepción,Concepción4030000,Chile*   
c*DepartamentodeCienciasdelaComputación,FacultaddeIngeniería,UniversidaddeConcepción,Concepción4030000,Chile*   
d*CentrodeInvestigacionesFarmacognósticasdelaFloraPanameña(CIFLORPAN),Apartado0824-00178,FacultaddeFarmacia,UniversidaddePanamá,Ciudadde Panamá,Panamá*   
e*CBio3Laboratory,SchoolofChemistry,UniversityofCostaRica,SanPedro,SanJosé,CostaRica*   
f*LaboratoryofComputationalToxicologyandBiologicalTestingLaboratory(LEBi),UniversityofCostaRica,SanPedro,SanJosé,CostaRica*   
g*ComputationalBiologyandChemistryResearchGroup,VicerrectoradodeInvestigación,UniversidadCatólicadeSantaMaría,Arequipa04000,Perú*   
h*DepartmentofPharmaceuticalProducts,FacultyofPharmacy,UniversidadeFederaldeMinasGerais,Av.AntônioCarlos,6627,BeloHorizonte,MinasGerais 31270-901,Brazil*   
i*AdvancedComputingLab(CNCA),NationalHighTechnologyCenter(CeNAT),Pavas,SanJosé,CostaRica*

|  |  |  |
| --- | --- | --- |
| article | info | abstract |
| *Keywords:*  Artificialintelligence Chemoinformatics  Drugdiscovery  Education  Openscience |  | Chemoinformaticsisanindependentinter-disciplinewithabroadimpactindrugdesignanddiscovery,medicinal chemistry,biochemistry,analyticalandorganicchemistry,naturalproducts,andseveralotherareasinchemistry. Throughcollaborations,scientificexchanges,andparticipationininternationalresearchnetworks,LatinAmer-icanscientistshavecontributedtothedevelopmentofthissubject.Theaimofthisperspectiveistodiscuss thestatusandprogressofthechemoinformaticdisciplineinLatinAmerica.Weteamuptoprovideanauthor´s perspectiveonthetopicsthathavebeeninvestigatedandpublishedoverthepasttwelveyears,collaborationsbe-tweenLatinAmericaresearchersandothersworldwide,contributionstoopen-accesschemoinformatictoolssuch aswebservers,andeducational-relatedresourcesandevents,suchasscientificconferences.Weconcludethat linkingandfosteringcollaborationwithineachnationaswellasamongotherLatinAmericannationsandglob-allyismadepossiblebyopenscienceandthedemocratizationofscience.Wealsooutlinestrategicactionsthat canboostthedevelopmentandpracticeofchemoinformaticintheregionandenhancetheinteractionbetween LatinAmericancountriesandtherestoftheworld. |

**1.Introduction**

Chemoinformaticsisanindependentdisciplinethathasabroad rangeofapplicationsinchemistry[1,2].Ithasseveralformaldef-initions,forexample,“*Allconceptsandmethodsthataredesignedto interfacetheoreticalandexperimentaleffortsinvolvingsmallmolecules*”[3]or“*Chemoinformaticsisagenerictermthatencompassesthedesign, creation,organization,management,retrieval,analysis,dissemination,vi-sualization,anduseofchemicalinformation*” [4].Otherdefinitionsare collectivelydiscussedelsewhere[1].Itshouldbenotedthatseveral termspublishedintheliteraturerefertothesamedisciplineandits

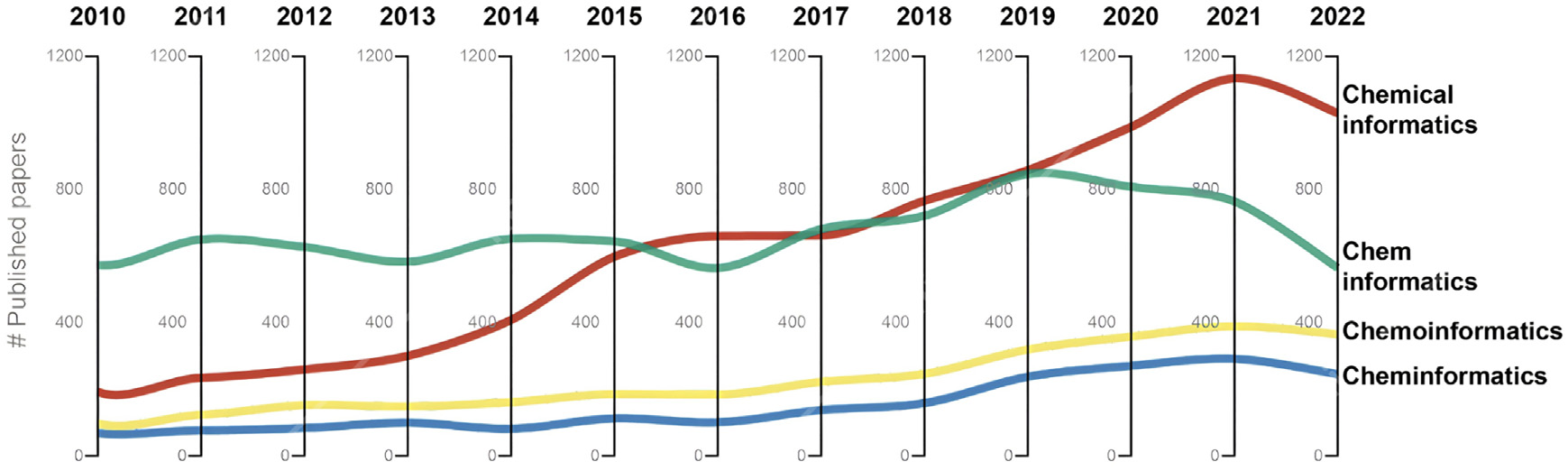
usesometimesdependsonthegeographicalregion.Toillustratethis fact,Fig.1showsaparallelcoordinateplotwiththenumberofpa-persinPubMedintheperiod2010–2022withthekeywords“Chemi-calinformatics”,“Cheminformatics”,“Chemoinformatics”,and“Chem informatics.” Thefigureshowsanoverallincreaseinthenumberof publicationssince2010withaslightdecreasein2022.Amongthetwo frequentlyusedtermsinconferencesandscientificjournals,“chemoin-formatics” showsaslightlylargernumberofpublicationsthanthose containingtheterm“cheminformatics”.Itshouldbeemphasizedthat thecommunityhasnotyetagreedonasingletitle,unliketheclosely relateddiscipline"bioinformatics,"whosetermhasmostlybeenap-

*Abbreviations:*AI,artificialintelligence;LaNAPDB,LatinAmericanNaturalProductDatabase;QSAR,quantitativestructure-activityrelationships;SAMPL,statis-ticalassessmentofthemodelingofproteinsandligands;VOS,visualizationofsimilarity.  
 ∗Corresponding[author.](mailto:medinajl@unam.mx)

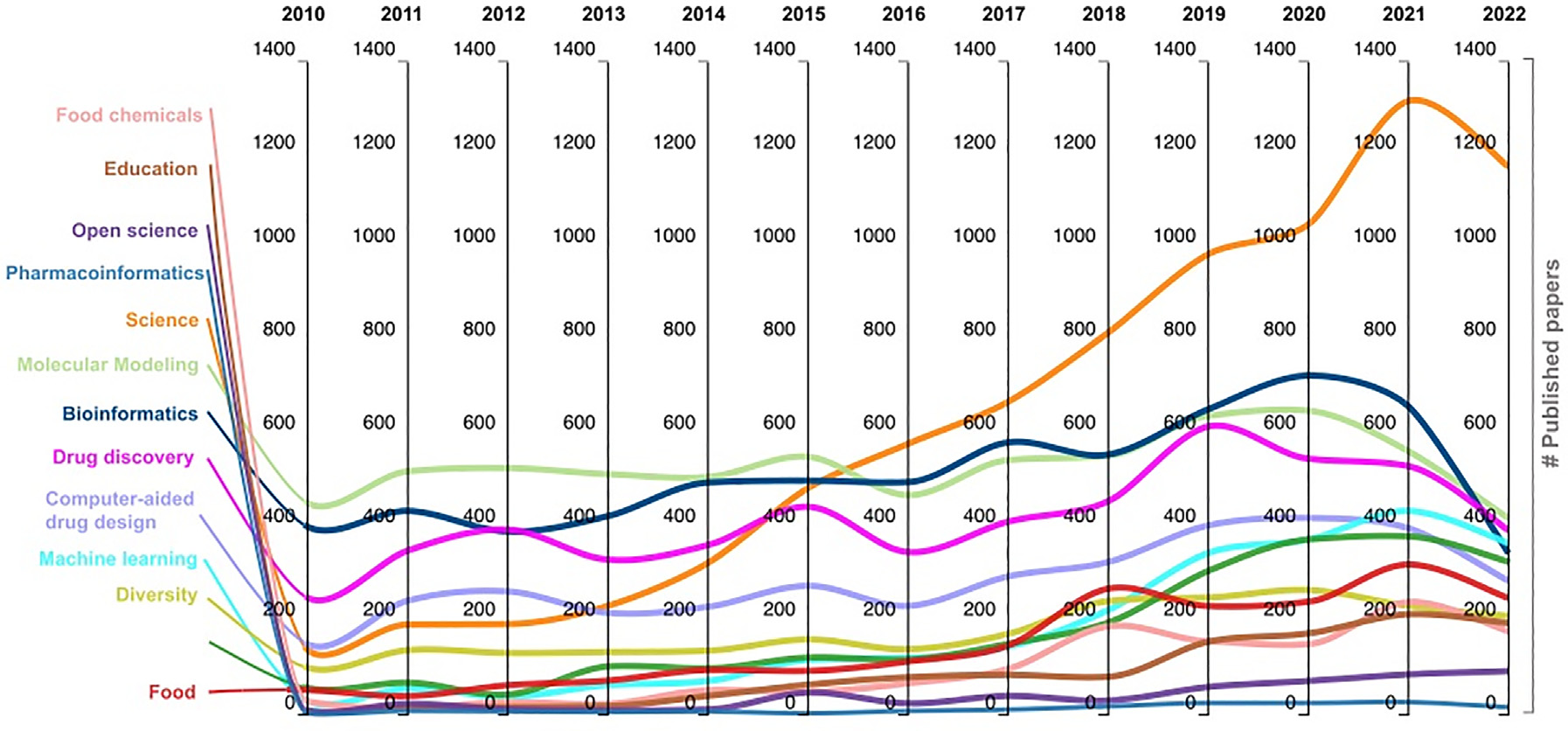
*E-mailaddress:*[medinajl@unam.mx](mailto:medinajl@unam.mx)(J.L.Medina-Franco).

<https://doi.org/10.1016/j.ailsci.2023.100077>  
[Received6May2023;Receivedinrevisedfor](https://doi.org/10.1016/j.ailsci.2023.100077)m10May2023;Accepted10May2023   
Availableonline11May2023   
2667-3185/© 2023TheAuthors.PublishedbyElsevierB.V.ThisisanopenaccessarticleundertheCCBYlicense(<http://creativecommons.org/licenses/by/4.0/>)

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.1.**NumberofpapersinPubMedintheperiod2010–2022withthekeywords“Chemicalinformatics”,“Cheminformatics”,“Chemoinformatics”,and“Chem informatics.” Thefigureshowsthetrendinusageofthemostcommontermsthatreferstothesamediscipline.

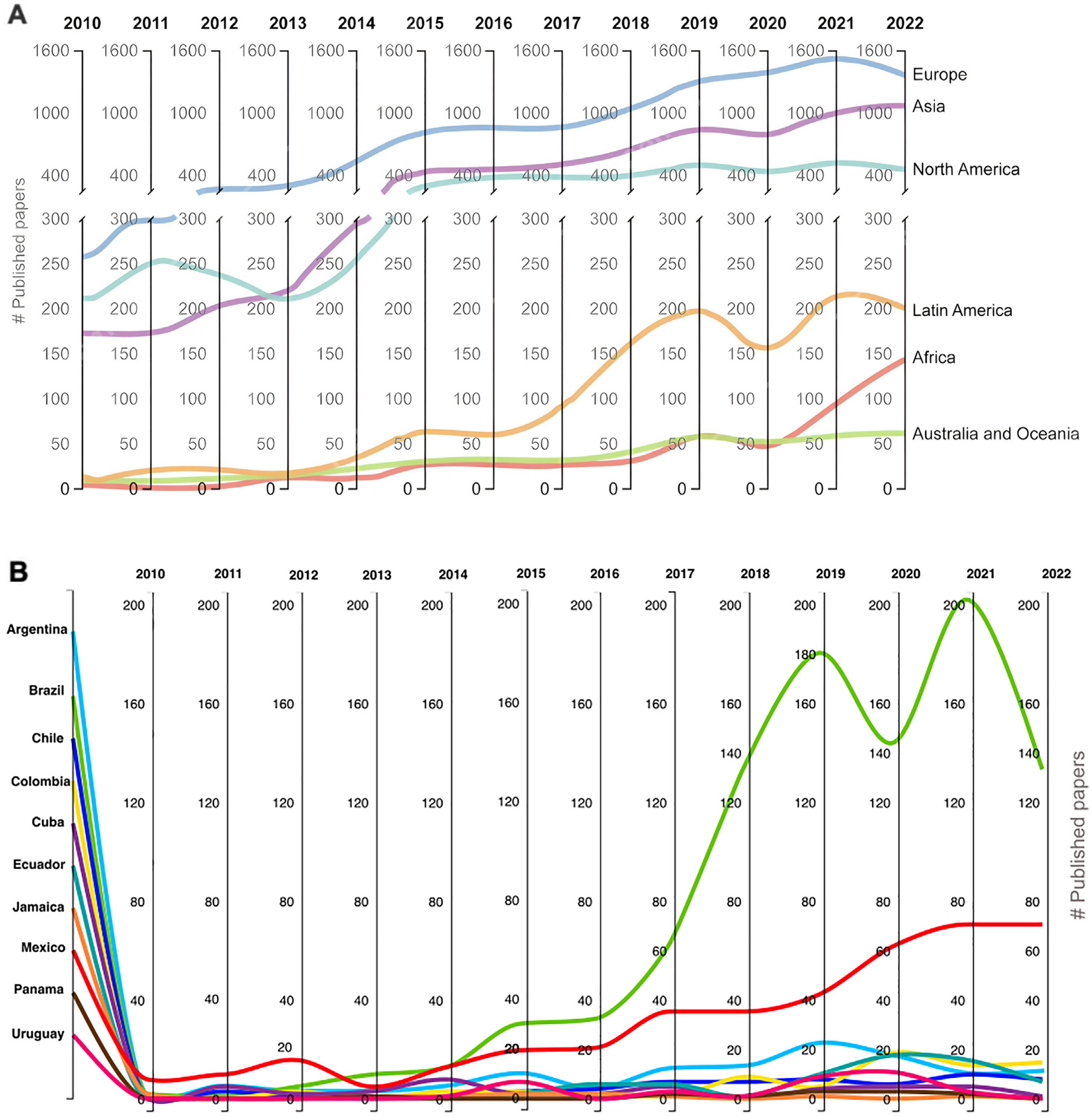


**Fig.2.**ThenumberofpapersinPubMedwiththekeywords“Chemicalinformatics”,“Cheminformatics”,“Chemoinformatics”,and“Cheminformatics” plusanother relevantkeywordbyyear.

proved[2].Fortheremainderofthismanuscript,wewillusetheterm“chemoinformatics.”  
 Regardlessofthespecificterm,sincethefirstapplicationsof chemoinformaticsinthe’50s,thedisciplinehasbeenevolvingandex-pandingrapidlywiththemostrecentandlargeadvancesinartificialin-telligence(AI)[5].Chemoinformaticshasabroadrangeofapplications acrossdifferentresearchareas,includingdrugdesign,chemogenomics, andsystemspharmacology,aswellasmedicinal,analytical,organic, cosmetic,food,andinorganicchemistry[6–8],andnaturalproductre-search[6,9],amongothers.Toillustratetheapplicationsofchemoin-formaticsinotherareas,Fig.2showsthenumberofpublicationsin PubMedwiththekeywords“Chemicalinformatics”,“Cheminformatics”,“Chemoinformatics”,and“Cheminformatics” plusanotherrelevantkey-wordbyyear.Theanalysisrevealsthatwhereasmolecularmodeling wastheprimarytopicconnectedtochemoinformaticsatthebeginning oftheexaminedperiod(2010–2022),otherfieldshavebeeninvolved withtime,witheducationbeingoneofthem.Thishasallowedusto glimpsethetremendousgrowthofthisdisciplineanditssignificantap-plicationinrelatedareasastheyearspassby.Inthesameperiod(2010–2022),themostfrequentareaspublishedalongwithchemoinformatics (asmeasuredbythenumberofpublicationswithconcurrentkeywords) weremainlymolecularmodelinganddrugdiscoverywith6935and 5459papers,respectively.Otherfrequentareaspublishedalongwith chemoinformaticswerebioinformatics(6689),computer-aideddrugde-

2

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.3.**NumberofpapersinPubMedintheperiod2010–2022withthekeywords“Chemicalinformatics”,“Cheminformatics”,“Chemoinformatics”,and“Chem

informatics” **A.**Analyzedbycontinent.AmericaisseparatedasNorthandLatinAmerica.**B.**AnalyzedbyeachLatinAmericancountry.

oratesonfuturedirections,challenges,andperspectivestostrengthen chemoinformaticsinLatinAmerica.Section6presentssummarycon-clusions.

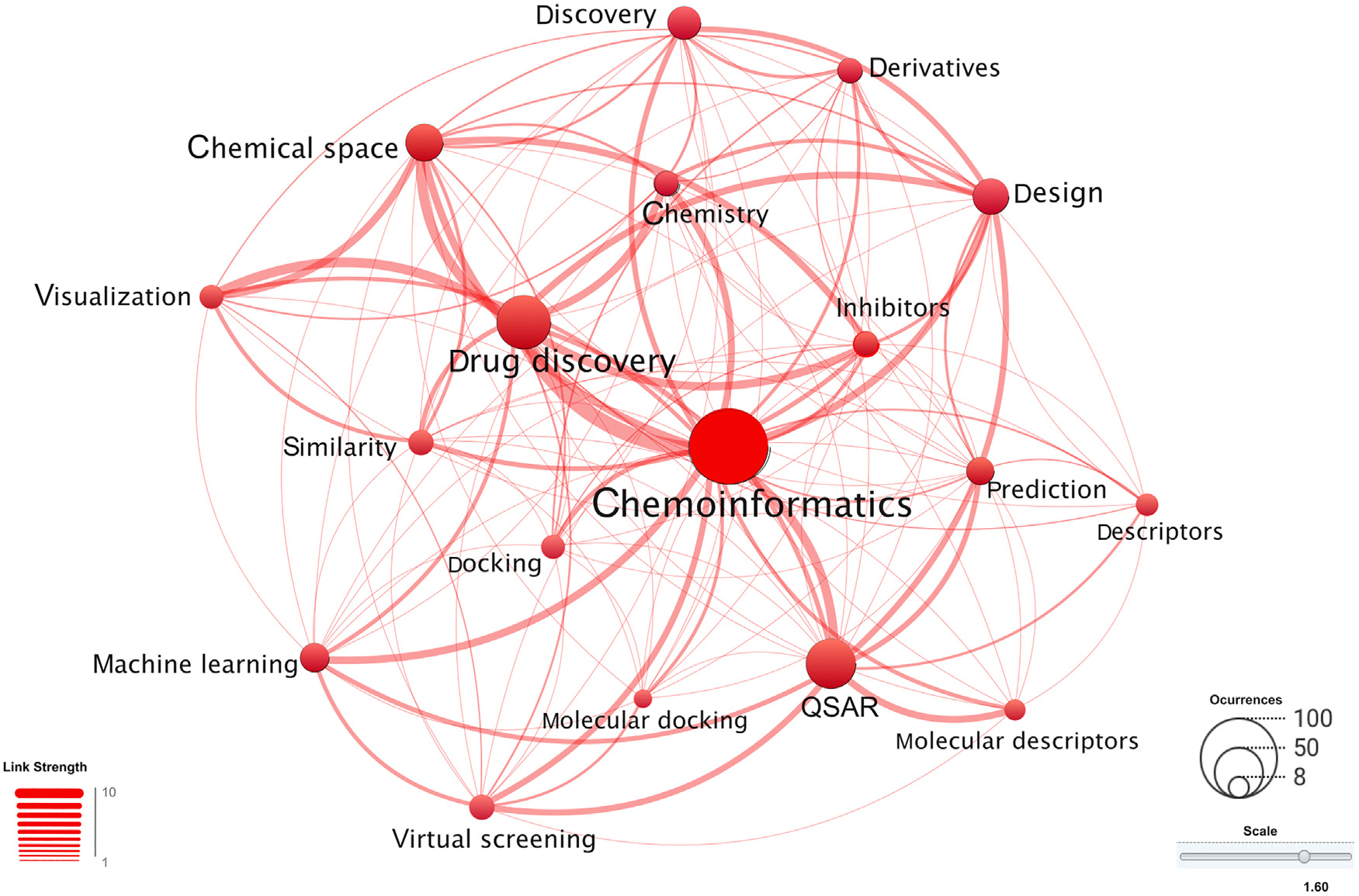
**2.Research:publications,applications,andcollaborations**

*2.1.Publicationsandmainapplications*

Inthissection,wediscusscontributionsbyresearchgroupsinLatin Americatowardsthedevelopmentandapplicationofchemoinformat-icsasanalyzedthroughpeer-reviewedpublicationsbetween2010and 2022.Fig.3depictsthetremendousexpansionofchemoinformaticsin LatinAmericaduringtheprevious12years.Fig.3Aunderlinesthat, duringthepreviousfiveyears,LatinAmerica’spublicationratehassig-nificantlyincreased,eventhoughEurope,Asia,andNorthAmericahave publishedmanymorepapersthanLatinAmericaoverall.Fig.3Bindi-catesthat,astheyearspassby,moreLatinAmericancountrieshavebeen joiningthedevelopmentofthisdiscipline,contributingtomorepub-lishedpeer-reviewedpapers.ByputtingtogetherthestudiesofallLatin

3

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.4.**BibliometricmapbasedonWebofSciencedatabasedonkeywordco-occurrenceduringtheperiod2010–2022.Thedatawereobtainedusingthekeywords

“Chemicalinformatics”,“Cheminformatics”,“Chemoinformatics”,and“Cheminformatics” andeachoftheLatinAmericancountrieswithpublishedpapersin2010–

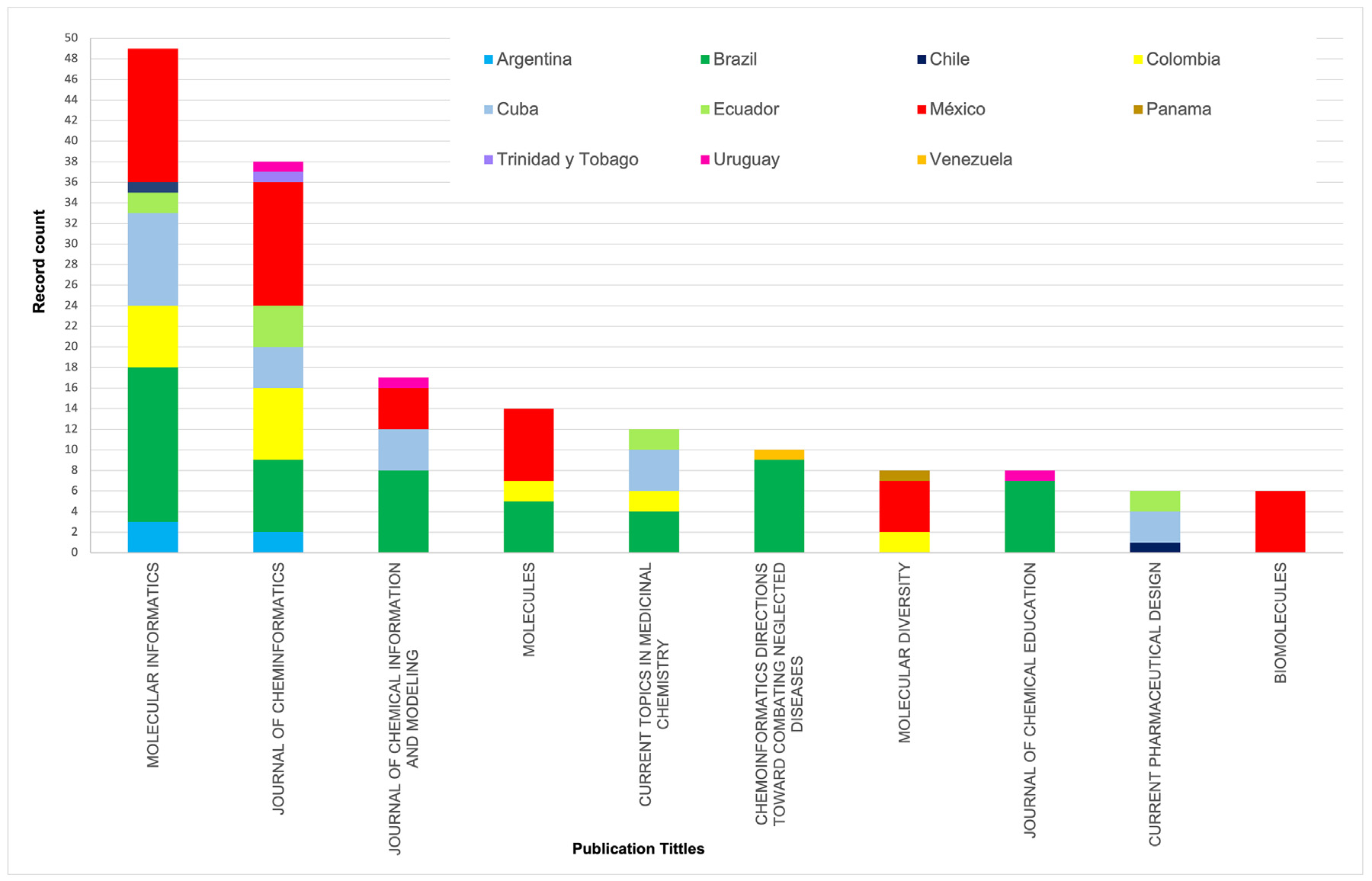
2022.Atotalof474publicationswereidentified.Therelativesizeofthenodesisrelatedtothenumberofpublicationscontainingthekeyword.

icsORChemicalinformatics” foreachoftheLatinAmericancountries. Surprisingly,therewasasignificantdifferencebetweenthedataob-tained,findingmoretotalpublicationsobtainedinWebofScience(536) incontrastwithPubMed(451),whichalsowerereflectedinfewerpeer-reviewedpublicationsfoundpercountryasitisdepictedinFig.S5. Fig.4showsanetworkwiththekeyconceptsrelatedtochemoinfor-maticsineachLatinAmericancountryasfoundintheWebofScience. ThenetworkwascreatedwithVisualizationofSimilarity(VOS)viewer desktopversion1.6.19(0)(<https://www.vosviewer.com/>)[11]based onkeywordconcurrentduri[ngtheperiod2010–2022.Th](https://www.vosviewer.com/)isfigurede-pictsthemainconceptsinvolvedinchemoinformatics.Eachkeyword anditsentanglementaregivenbytheusageofpublications,anditis alsonotablethethicknessoftheedgesinthegraph,wherechemoinfor-maticsisstronglyassociatedwithdrugdiscoveryandchemicalspace. Thenetworkshowsthatdrugdiscovery,quantitativestructure-activity relationships(QSAR),andchemicalspaceareamongstthemostfrequent con-currentwordsandapplicationspublishedsofarbyLatinAmerican scientists.

Fig.S2(Supplementarydata)highlightstheprincipalareasofap-plicationinwhichthepublicationsofeachLatinAmericancountry arefocusedaccordingtotheWebofSciencecategories.Theanaly-sisindicatesthatthemaincategoriesare“PharmacologyPharmacy”,“ChemistryMedicinal”,“ChemistryMultidisciplinary”,“ComputerSci-enceInformationSystems”,“ComputerScienceInterdisciplinaryAppli-cations”,BiochemistryMolecularBiology”,“ComputerScienceInfor-mationSystems”,“Neurosciences”,“BiochemistryMolecularBiology”, andMicrobiology”.Thecategory“ChemistryMultidisciplinary” pre-dominatesinBrazil,Mexico,Colombia,Ecuador,andCuba.ForChile andSaintLucia“PharmacologyPharmacy” isthemajorcategory.In Argentina,“ChemistryMedicinal” and“ComputerScienceInformation Systems” are,equally,themostfrequentcategories.Greatervariability offocusareasisevidentwithrespecttothepreviouslymentionedcoun-tries,andtheyexperiencedifferentiationbythepercentagesassigned toeachfocusarea.Likewise,therecentgrowthoffocusareassuchas

4

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.5.**Toptenpeer-reviewedjournalspublishingchemoinformaticspapersbyLatinAmericancountriesduring2010–2022accordingtotheWebofScienceCore

Collection.

Wecanfindagreatercollaborationwithrespecttothenumberofar-ticleswithNorthAmerica,butEuropealsogeneratesarelevantcollab-orationwithLatinAmericabyinvolvingalargernumberofcountries implicatedinsaidcollaborations(sixoutoftencountries).Finally,Asia andOceania,althoughtheyareonlyshownwithonecountrywithinthe toptencollaboratorsforeachofthesetwocontinents,theyillustratea relevantconnectionwithLatinAmericancountries.

RegionalcollaborationsareshowninFig.S3B,wherewevisualize theimportantsupportnetworkthatinvolvestheentireregion,andthat hasallowedLatinAmericatogrowinthedevelopmentofchemoinfor-matics.

Figs.S3CandS3Dshowtheinteractionsusinganalternativegraphic method,tofurtherhighlighteachcollaborationbetweenLatinAmerican countriesandotherregionsoftheworld.Fig.S3CdistinguishesBrazilas theLatinAmericanscientificsuperpowerthathasthehighestnormal-izedcollaborationwiththedifferentcountriesdescribed,foregrounding amongthemUSA,Germany,Sweden,andJapan.Mexicoisthesecond countrywiththehighestnumberofcollaborationsbetweenthesame countriesandonlycomparesBrazilwithSpainregardingthenumberof collaborations.

Sincedatanormalizationwithrespecttothenumberofcollabora-tionsfromthetwoLatinAmericancountrieswiththehighestnumber ofpublicationsmakesitdifficulttovisualizethislastgraphfortherest oftheLatinAmericancountries,Fig.S3Dwasgenerated.Fig.S3Dal-lowsustoobserveingreaterdetailthedifferentiationbyeachLatin Americancountrywithrespecttoitscollaborationwiththerestofthe tencountrieswiththegreatestcollaboration,visualizingthatArgentina, Chile,Colombia,andPeruingeneralhaveverysimilartrends,differ-inginthatPeruhasaverymarkedandgreatercollaborationwithItaly whileArgentina,ChileandColombiahaveitwithSpainindifferent proportions.

SimilartrendsareperceivedwiththerestoftheLatinAmerican countries,whereEcuador,likeBrazil,hasagreatercollaborationwith theUSA.Jamaica,Uruguay,Panama,Cuba,Venezuela,andCostaRica remaininalimitedintervalofcollaborations.Finally,Nicaragua,Barba-

5

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*

**Table1**

Exemplarywebservers,standalonetools,andotherresourcesdevelopedinLatinAmerica.

|  |  |  |  |
| --- | --- | --- | --- |
| Country | Resource | Application-use | Ref.,URL |
| Argentina | LIDeBTools | Publiclyavailableopen-sourcecustomizable  cheminformaticstoolstobeusedincomputer-assisteddrug discovery.  MolecularDynamicsSimulationsandcustomsoftware toolsforanalyzingthestructureandthermodynamicsof solutionsusingsimulationdata.  Computationalpolypharmacologywithwetlabanalysesto strengthenthedrugdesignanddevelopmentprocesses.  Morethanresourcesdoneandunderdevelopmentfor modelingandanalyzingproteinsequencesandstructures. Computersimulationsatdifferentlevelsofresolutionand complexitytodecipherhowbiologicalmacromolecules interactwitheachothertoaccomplishtheirfunction. CodesinGitHubrepositoriesforthecalculationofuseful descriptorsinchemoinformatics:i)pH-dependent  lipophilicdescriptorsforaminoacidsii)descriptorsbased onthecountofnconjugatedsystemswith*n*=2(*n*=*𝜋* electrons/2)relevantforthepredictionofpartition  coefficientsinanaromaticsolventsuchastoluene.  Fivefreewebservicestoanalyzechemicaldiversityand chemicalspace,generate*insilico*librariesofpeptides,and predictbioactivityprofiles.  Dedicatedtothedevelopmentofmethodsthatallow advancedsimulationstobecarriedoutbutwithlow computationalcost,toimprovethecomparabilityof theoreticalstudieswithbiochemical,biophysical,or molecularbiologyexperiments. | [34],[https://lideb.biol.unlp.edu.ar/?p=1237](https://www.lideb.biol.unlp.edu.ar/?p=1237) |
| Brazil | UniCamp | <http://leandro.iqm.unicamp.br/m3g/main/home.shtml> |
| Chile | Pharmacoinformaticsanddrug design  CBSMCenterofBioinformatics andMolecularModeling  MPTG-CBP | <https://ramirezlab.github.io/5.1_resources> |
| Chile | <https://cbsm.utalca.cl/?page_id=2073> |
| Colombia | <https://mptg-cbp.github.io/> |
| CostaRica | Cbio3Laboratory | <https://cbio3group.netlify.app/> |
| Mexico | DIFACQUIM,d-Tools | <https://www.difacquim.com/d-tools/> |
| Uruguay | InstitutPasteurdeMontevideo | [https:](https://pasteur.uy/laboratorios/simulaciones-biomoleculares/)  [//pasteur.uy/laboratorios/simulaciones-biomoleculares/](https://pasteur.uy/laboratorios/simulaciones-biomoleculares/) |

putationalstudiesbeyondchemoinformaticse.g.,toolsassociatedwith pharmacoinformaticsingeneral,moleculardynamics,andbiophysics. Withinchemoinformatics,thewebserversandtoolsaremostlytuned fordrugdiscoveryanddevelopmentofleadcandidatesbutalsocanbe usedforotherresearchareas,asdiscussedinSection2.1.

*3.1.Compounddatabases*

Concerningthedevelopmentofcompounddatabases,perhapsoneof themostnotablecontributionsintheregionisthegeneration,curation, andupdateofnaturalproductsdatabasesfreelyaccessible.Formore thantenyears,Brazilhasbeendevelopinganextensivedatabaseofnat-uralproducts,andseveralotherLatinAmericancountrieshavefollowed theinitiative.Table2summarizesexemplarynaturalproductdatabases inthepublicdomain.Thestatusofthedevelopmentofnaturalproduct databasesinLatinAmericahasbeenreviewedrecently.Notably,thereis amulti-countryefforttoputtogetheraunifiedLatinAmericanNatural ProductDatabase– LANaPDB[15,16].

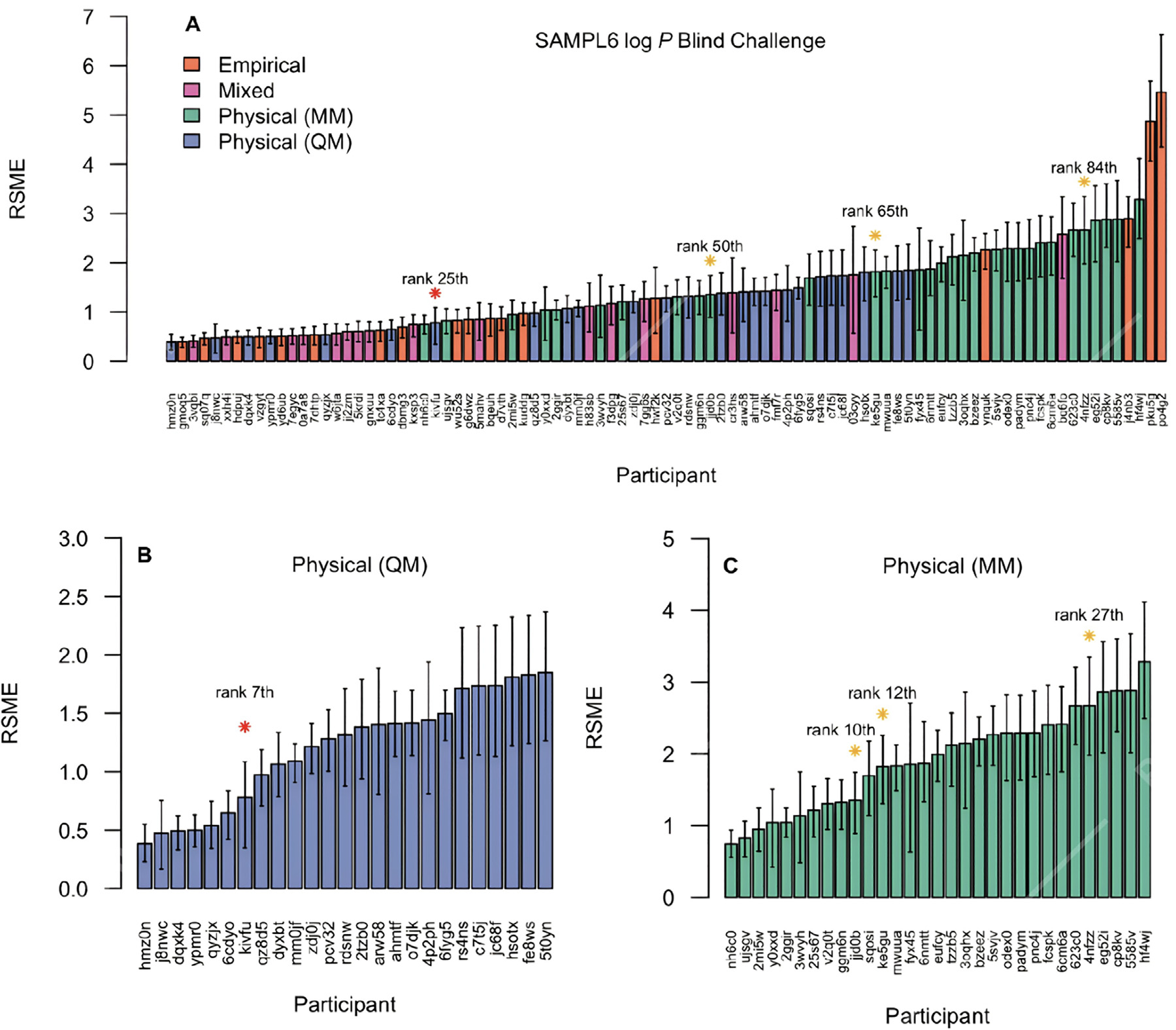
*3.2.Blindchallengesforvalidatingchemoinformaticsmethods*

Asmentionedintheprevioussection,theLatinAmericanregion, throughitsresearchgroups,hasgeneratedknowledgeregardingthede-velopmentofdatabases,webservers,andchemoinformaticstools.Ina complementaryway,theblindchallengesknownasStatisticalAssess-mentoftheModelingofProteinsandLigands(SAMPLChallenges[17]) fundedbytheUSANationalInstitutesofHealthhasprovidedanexcel-lentopportunityforLatinAmericanstovalidatetheimpactoftheirtools andmethodologiesthroughtheparticipationofresearchersinthese challenges.Thesectiononthepredictionofphysicochemicalproper-tiesinSAMPLChallengeshasincludedpredictionsofsolvationenergies [18,19],acidityconstant[20,21]andpartitioncoefficients[21,22]in variousbiphasicsystems(cyclohexane/waterand*n*-octanol/water)rel-evantforrationaldrugdesign.

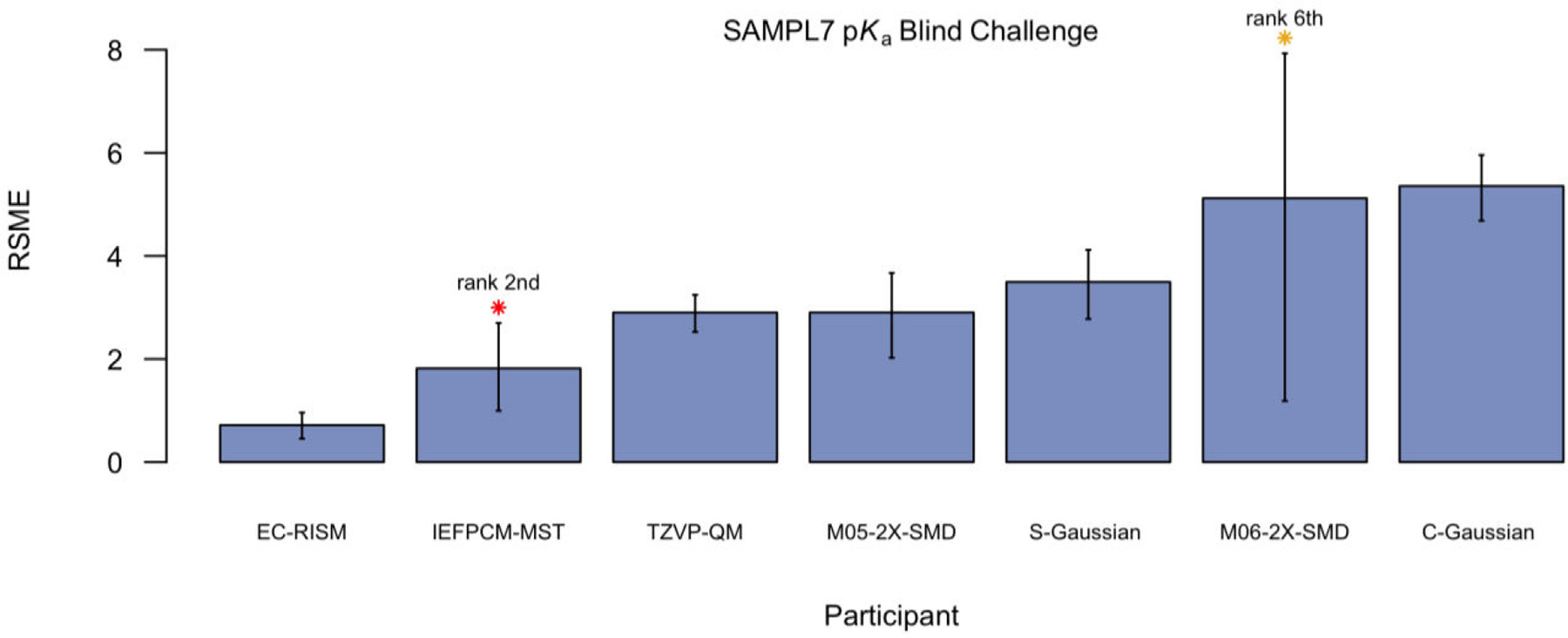
Table3showstheparticipationofLatinAmericanresearchgroupsin predictingphysicochemicalpropertiesofdrug-likemoleculesinthepe-riodanalyzed(2010–2022).Itcanbenotedthatourparticipationinthe SAMPL-3,SAMPL-4,andSAMPL-5blindchallengeswasnull,however,

6

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.6.**Representationoftheroot-meansquareerror(RMSE)of**A.**the91submissionstotheSAMPL6*n*-octanol/waterlog*P*predictionchallenge(theparticipation ofgroupsfromLatinAmericaaremarkedwithstarswheretheredstarrepresenttothesubmissionoftheUniversityofCostaRicaincollaborationwiththeUniversity ofBarcelonaandyellowstarsthesubmissionsoftheAustralUniversidadofChileincollaborationwiththeUniversityofConcepción),**B.**Twentyfivesubmissions providedbyphysicalmethodsusingquantummechanicalapproaches,and**C.**Thirtyonesubmissionsprovidedbyphysicalmethodsusingmolecularmechanics approaches.TherankobtainedbyeachLatinAmericansubmissionbothgloballyandbycategoryisrepresentedatthetopofeachmark.Detailsofthemethodsand participantscanbefoundinTableS1(Supplementarymaterial)and<https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties>.



**Fig.7.**Representationoftheroot-meansquareerror(RMSE)ofsevensubmissionstotheSAMPL7tothep*K*achallenge.TheparticipationofgroupsfromLatin AmericaaremarkedwithstarswheretheredstarrepresentsthesubmissionoftheUniversityofCostaRicaincollaborationwiththeUniversityofBarcelonaand theorangestarthesubmissionsoftheNationalUniversityofSantiagodelEstero,ArgentinaincollaborationwiththeMiamiUniversity.Therankobtainedby eachLatinAmericansubmissionisrepresentednexttoeachmark.DetailsofthemethodsandparticipantscanbefoundinTableS2(Supplementarymaterial)and <https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/pKa>.

7

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*

**Table2**   
ExemplarycompounddatabasesofnaturalproductsgeneratedinLatinAmerica.

|  |  |  |  |
| --- | --- | --- | --- |
| Database | Country | Generaldescription(link)a | Refs. |
| NaturAr  NuBBEDB | Argentina Brazil | NaturaproductsfromArgentina.<https://naturar.quimica.unlp.edu.ar/es/> NaturalproductsofBrazilianbio[diversity.DevelopedbytheSãoPauloSt](https://naturar.quimica.unlp.edu.ar/es/)ateUniversityandtheUniversityofSão [Paulo.Includesnaturalproductsfromplants,micro](http://nubbe.iq.unesp.br/portal/nubbe-search.html)organisms,terrestrialanimals,andmarineanimals.  <http://nubbe.iq.unesp.br/portal/nubbe-search.html> [Databasecomposedofsecondarymetabolitesfromp](http://nubbe.iq.unesp.br/portal/nubbe-search.html)lants,developedattheFederalUniversityofParaiba.  <https://sistematx.ufpb.br/> [Naturalp](https://sistematx.ufpb.br/)[roductsfromplantsthathavebeenpublish](http://zinc12.docking.org/catalogs/uefsnp)edseparately.DevelopedattheStateUniversityofFeirade Santana.<http://zinc12.docking.org/catalogs/uefsnp> Brazilian[CompoundLibraryisanewmanuallycura](http://zinc12.docking.org/catalogs/uefsnp)tedvirtualsubstancelibrarydevelopedbyBrazilianresearch groupstosupportcomputer-aideddrug*design*work.ThefirstversionofBraCoLihas1176substances.Contains biologicalandchemicalinformationonsynthetic,semi-syntheticandnaturalsubstances.  AntiviralMedicinalPlantsandNaturalProductsisaDBdevelopedinBrazilandcontainsbioactivesubstances frombiodiversitywithantiviralactivity.  Naturalproductsisolatedfromplantsandmicroorganisms.ThedatabaseisbeingdevelopedbytheCbio3 Laboratory,SchoolofChemistryattheUniversityofCostaRica.  DevelopedbytheResearchLaboratoryinNaturalProductsoftheUniversityofElSalvador.  Naturalproductsfromplantsmainl[yisolatedandcharacterizedattheIns](https://uniiquim.iquimica.unam.mx/)tituteofChemistryoftheNational AutonomousUniversityofMexico.<https://uniiquim.iquimica.unam.mx/> Naturalproductsisolatedandchar[acterizedinMexicoattheSchoolofC](https://uniiquim.iquimica.unam.mx/)hemistryoftheNationalAutonomous UniversityofMexi[coandotherMexicaninstitutions.Includescompoundsfromplants,f](https://figshare.com/articles/dataset/BIOFAQUIM_V2_sdf/11312702)ungi,propolis,andmarine animals.version2[https://figshare.com/articles/dataset/BIOFAQUI](https://figshare.com/articles/dataset/BIOFAQUIM_V2_sdf/11312702)M[\_V2\_sdf/11312702](https://figshare.com/articles/dataset/BIOFAQUIM_V2_sdf/11312702)  Naturalproductsf[romplantsthathavebeentestedinover25invitroandinvivobioass](https://figshare.com/articles/dataset/BIOFAQUIM_V2_sdf/11312702)aysfordifferent therapeutictargets.DevelopedattheUniversityofPanama.  [DatabasecreatedattheCa](https://perunpdb.com.pe/)tholicUniversityofSantaMariawithnaturalproductsfromanimalsandplants.  <https://perunpdb.com.pe/> [Freeaccessibledatabaseco](https://perunpdb.com.pe/)ntainingthe3Dstruc[turesofseveralplantsecondarycompoundsalongwith](http://pscdb.appsbio.utalca.cl/viewSearch/index.php)their physicochemicalandpharmaceuticalproperties.<http://pscdb.appsbio.utalca.cl/viewSearch/index.php> | [35]  [36,37] |
| SistematX | Brazil | [38,39] |
| UEFS | Brazil | [40] |
| BRACOLI | Brazil | [41] |
| avMpNp | Brazil |
| NAPRORE-CR | CostaRica |
| LAIPNUDELSAV UNIIQUIM | ElSalvador Mexico | [42] |
| BIOFACQUIM | Mexico | [43,44] |
| CIFPMA | Panama | [45] |
| PeruNPDB | Peru | [46] |
| PSC-db | Chile | [47] |
| aWhenavailable. |  |  |  |

**Table3**   
ParticipationofLatinAmericanresearchgroupsinglobalchallengesfortheStatisticalAssessmentoftheModelingofProteinsandLigands(SAMPLChallenges[17]).

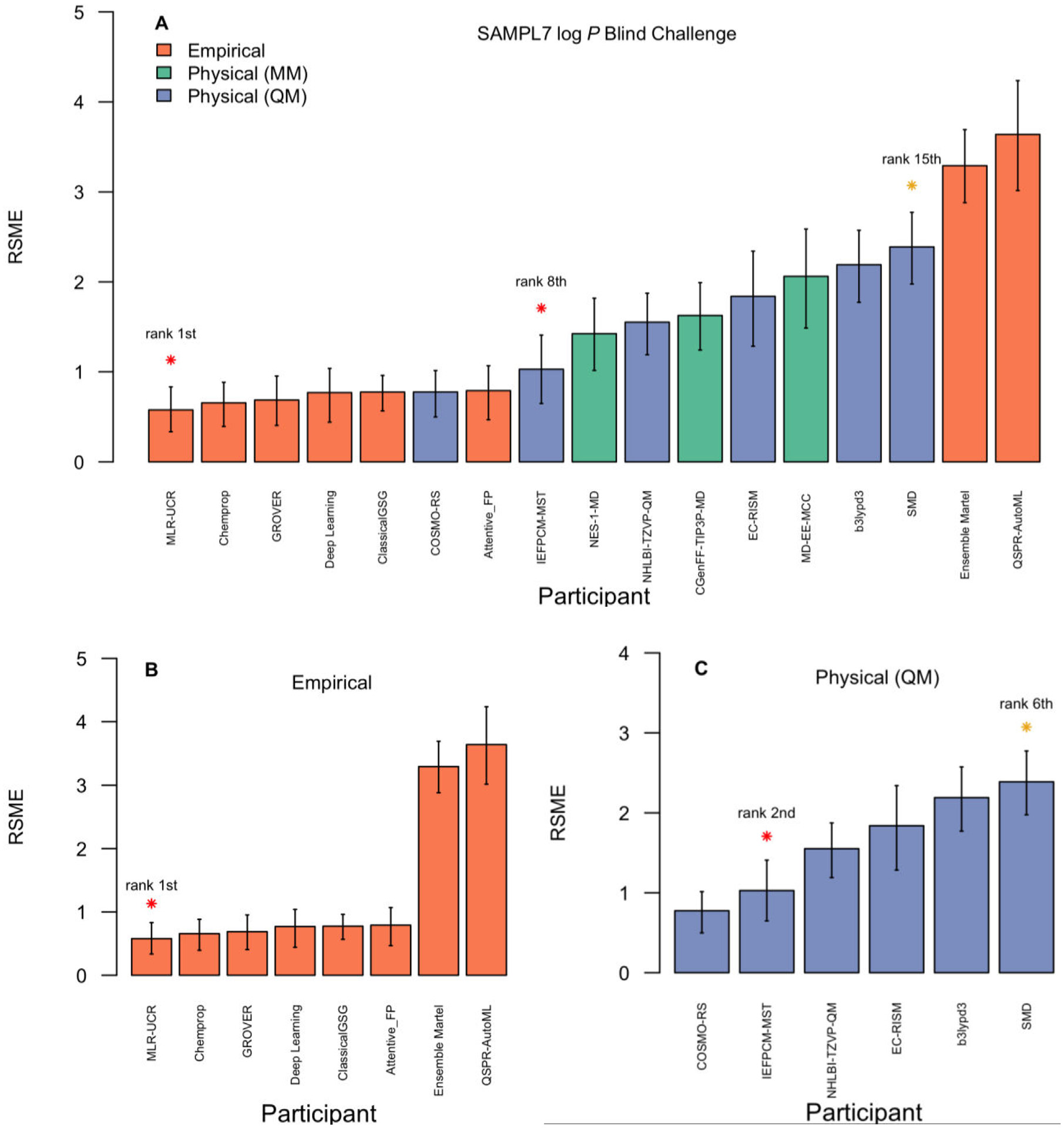
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Challenge | Physicochemicalblindchallenge | Participants | LatinAmericanparticipants inspecialissuepublications | Information | Refs. |
| **SAMPL3**  (2011–2012) | Predictionof36solvationfreeenergies forthreeseriesofchlorinatedcompounds | 21submissions  8different  groups/participants 49submissions  19different  groups/participants 76submissions  18different  groups/participants 37submissions  11different  groups/participants 91submissions  27different  groups/participants 9submissions  7different  groups/participants 33submissions  17different  groups/participants 6submissions  6different  groups/participants | 0 | [https://github.com/samplchallenges/ SAMPL3/tree/main/solvation\_energy](https://github.com/samplchallenges/SAMPL3/tree/main/solvation_energy) | [18] |
| **SAMPL4**  (2013–2014) | Predictionofhydrationfreeenergiesfora seriesof47smallmolecules. | 0 | [https://github.com/samplchallenges/ SAMPL4/tree/main/](https://github.com/samplchallenges/SAMPL4/tree/main/hydration_free_energy)  [hydration\_free\_energy](https://github.com/samplchallenges/SAMPL4/tree/main/hydration_free_energy)  [https://github.com/samplchallenges/](https://github.com/samplchallenges/SAMPL4/tree/main/hydration_free_energy) [SAMPL5/tree/main/](https://github.com/samplchallenges/SAMPL5/tree/main/distribution_coefficients)  [distribution\_coefficien](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/pKa)ts  [https://github.com/samplchallenges/ SAMPL6/tree/master/](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/pKa)  [physical\_properties/pKa](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/pKa)  [https://github.com/samplchallenges/](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/pKa) [SAMPL6/tree/master/](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/logP)  [physical\_properties/logP](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/logP)  [https://github.com/samplchallenges/](https://github.com/samplchallenges/SAMPL6/tree/master/physical_properties/logP) [SAMPL7/tree/master/](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/pKa)  [physical\_property/pKa](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logP)  [https://github.com/samplchallenges/ SAMPL7/tree/master/](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logP)  [physical\_property/logP](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logP)  [https://github.com/samplchallenges/](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logP) [SAMPL7/tree/master/](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logD)  [physical\_property/logD](https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logD) | [19] |
| **SAMPL5**  (2015–2016) | Predictionofdistributioncoefficient  predictionswereintroducedforthefirst timeforasetof53smallmolecules  Predictionofmicroscopicand  macroscopicpKasof24smallorganic moleculesproteinkinaseinhibitors  Predictionoftheoctanol/waterpartition coefficients(logP)of11smallmolecules proteinkinaseinhibitors  PredictionofthepKafor  22N-sulfonamides | 0 | [50] |
| **SAMPL6**  (2017–2018) | 0 | [20] |
| **SAMPL6**  (2017–2018) | 4submissions  2groups(ChileandCosta Rica/Spain)  2rankedsubmissions  2groups(Argentinaand  CostaRica/Spain)  3groups(Argentina/USA, CostaRica/Spain,andCosta Rica)  2rankedsubmissions  2groups(Argentina/USA andCostaRica/Spain) | [22,51,52] |
| **SAMPL7**  (2019–2020) | [21,53,54,55] |
| **SAMPL7**  (2019–2020) | Predictionoftheoctanol-waterpartition coefficientsfor22N-sulfonamides |
| **SAMPL7**  (2019–2020) | Predictionoftheoctanol-water  distributioncoefficientsfor  22N-sulfonamidescombiningthepKa andlogPpredictions |

ofCostaRicaandtheUniversityofBarcelonausingtheIEFPCM-MST modelpresentedverygoodresultswithbothglobally(rank8th)and physical(rank2nd)methods.Consideringthedifficultyofthesechal-lenges,thecollaborationbetweentheNationalUniversityofSantiago delEstero,ArgentinawiththeMiamiUniversityhadalsogoodresults (rank15th).

Fig.9showstheperformancebasedontheroot-meansquareerrorof sixsubmissionstotheSAMPL7*n*-octanol/waterlog*D*predictionchal-lenge.Onlyquantumchemistry-basedmethodsconsideredtheeffectof solvent.Theconformationalspaceofthe22moleculeswasexplored usingcheminformaticstools.Asexpected,duetotheoutstandingp*K*a

8

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.8.**Representationoftheroot-meansquareerror(RMSE)of**A.**the17submissionstotheSAMPL7*n*-octanol/waterlog*P*predictionchallenge(theparticipationof groupsfromLatinAmericaaremarkedwithstarswheretheredstarrepresenttothesubmissionoftheUniversityofCostaRica(rank1st),UniversityofCostaRicain collaborationwiththeUniversityofBarcelona(rank8th)andorangestarsthesubmissionsoftheNationalUniversityofSantiagodelEstero,Argentinaincollaboration withtheMiamiUniversity(rank15th)),**B.**Eightsubmissionsprovidedbyempiricalmethodsusingchemoinformaticsandmachinelearningapproaches,and**C.**Six submissionsprovidedbyphysicalmethodsusingquantummechanicsapproaches.DetailsofthemethodsandparticipantscanbefoundinTableS3(Supplementary material)and<https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logP>.

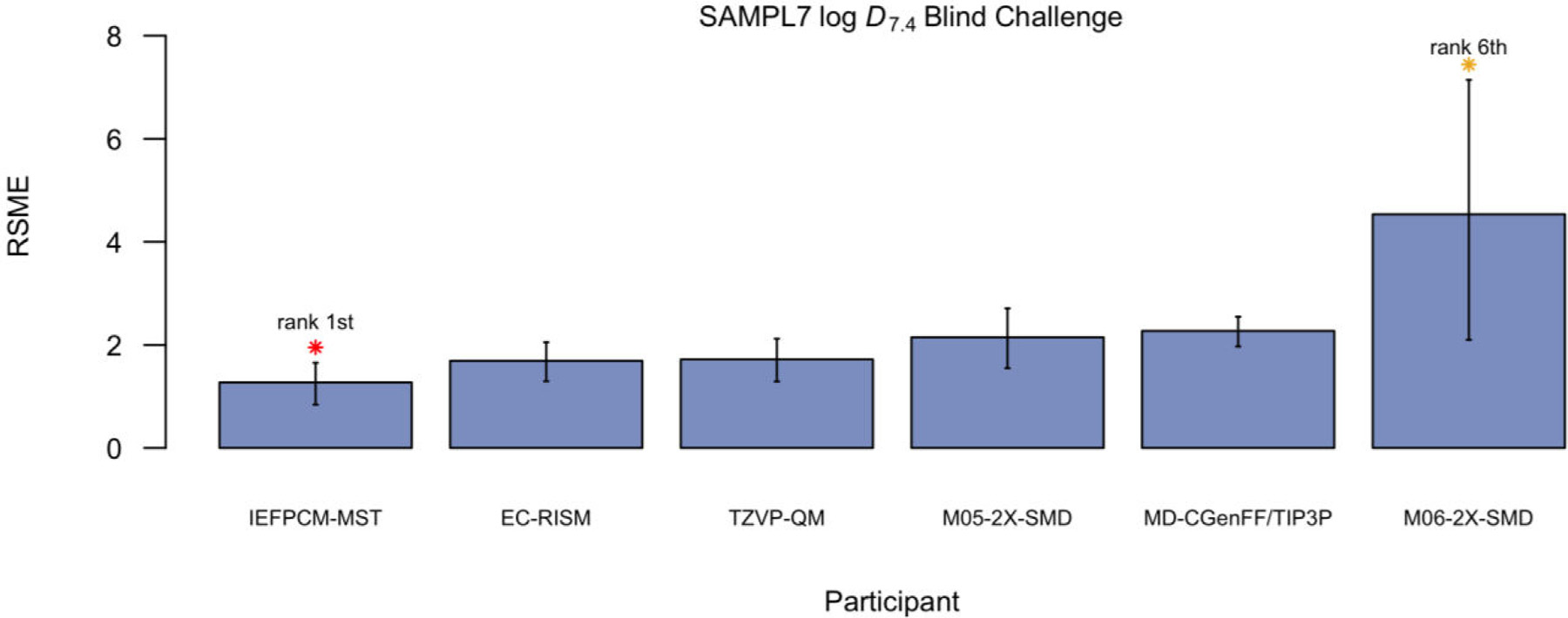
orationsbetweenacademicgroups,notonlywiththepharmaceutical industry,butalsofromthefoodandchemistryindustry,andinenviron-mentalpolicieswherethesepropertiesareusedtodetermineriskfactors forexposuretochemicals.

**4.Education,training,anddisseminationofscience**

Incontrasttobioinformaticswhereitiscommontofindspecial-izedcourses,textbooks,andeducationalresources,itisstillraretofind courses,andevenmore,graduateprogramsfocusedonchemoinformat-ics.ThisisnotonlyaparticularcaseofLatinAmericabutworldwide. Indeed,ascommentedelsewhere[2],chemistrydepartmentsandthe schoolsofpharmaciesinuniversitiesareconservativeinimplement-

9

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*



**Fig.9.**Representationoftheroot-meansquareerror(RMSE)ofsixsubmissionstotheSAMPL7*n*-octanol/waterlog*D*predictionchallenge.Theparticipationof groupsfromLatinAmericaaremarkedwithstarswheretheredstarrepresentsthesubmissionoftheUniversityofCostaRicaincollaborationwiththeUniversity ofBarcelonaandtheorangestarthesubmissionsoftheNationalUniversityofSantiagodelEstero,ArgentinaincollaborationwiththeMiamiUniversity.Therank obtainedbyeachLatinAmericansubmissionisrepresentednexttoeachmark.DetailsofthemethodsandparticipantscanbefoundinTableS4(Supplementary material)and<https://github.com/samplchallenges/SAMPL7/tree/master/physical_property/logD>.

ciatedwiththetheme(Fig.3Band3C).Inanevaluationofthelast20 years,thesepublicationsemphasizedtheBrazilianpanoramaregarding medicinalchemistry,naturalproductschemistryandmolecularbiology, addressingtheiradvancesandchallenges,especiallyinthesphereofed-ucationandresearch.Severalrelevanttopicsandhistoricallandmarks werecontextualizedanddiscussedthroughoutthisperiod,suchasthe roleofpharmaceuticalandchemistryeducationalinstitutionsinthede-velopmentofthearea,theintegrationofmultidisciplinaryknowledge requiredforitsexpansion,andtheevolutionintheareaforapplication ofsystemsandcomputationaltoolsdesignedtoextractknowledgefrom ahugesetofinformation[24,25].

Inthe1980s,thesectorwasdrivenbytheHumanGenomeProject, withgreatprogressintheareasofgenomicsandmolecularbiology, whoseessentialtechnologywasbioinformaticsforincreasinganalysis ofdatageneratedinDNA,RNAandproteinsequences[26].Therefore, bythelate1990stherewereasignificantnumberofbiomoleculesin GenBank.Also,inthe90′stheuseofQSARspreadworldwide.InBrazil duetointernationalcollaborations,especiallywithProfessorHugoKu-binyi,fromtheUniversityofHeidelberg,Germany,and,consecutively, coursesandworkshopsgiveninthecountry.Oneofthepoliticalconse-quencesofthiseventwasthecreationoftheChemicalStructureandBi-ologicalActivityDivisionintheBrazilianChemicalSociety,forerunner oftheMedicinalChemistryDivision,thelatterformedin1998[27,28]. Alreadyinthe2000s,thediscussionsthatpermeatedchemoinfor-maticsinthescientificandeducationalcontextwerebasedonthestan-dardizationoftermsandmethods,suchastheelaborationofaglossary oftermsusedindrugplanningandtheimprovementoftools,partic-ularlydockingandmoleculardynamics[29–31].Theclassicalversus industrialparadigmintheprocessofdrugdiscoverywasalsoadebated issue,aswellastheavailabilityofalargevirtualdatabaseofsmall moleculesandproteins(PDB)ofnaturalandsyntheticorigin,drovethe developmentofsoftwarefor*insilico*screening,appliedtotheapproach knownasvirtualscreening[32].

Onthisoccasion,thepublishedbooksauthoredbydistinguished Brazilianprofessorsandresearchers,suchas"MedicinalChemistry:the molecularbasisofdrugaction"byBarreiro&Fraga,"MethodsofTheo-reticalandMolecularChemistry"byMorgon,and"Bioinformatics:from biologytomolecularflexibility"byVerli,becameindispensablebiblio-graphicreferencesindisciplinesfocusedonmedicinalchemistryinun-dergraduateandgraduatecoursesthroughoutBrazil.Furthermore,the developmentofnew,saferandmoreeffectivedrugsformanydiseases hasbeendemonstrated,aswellastheinfluenceofeducationalinstitu-

10

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*

**Table4**

Examplesofrecentconferences,workshops,andeducationalresourcesgeneratedinLatinAmericawithafocusonchemoinformaticsordirectlyassociateddisciplines.

|  |  |  |  |
| --- | --- | --- | --- |
| Type | Eventorresource | Description(targetaudience) | Access/Refs. |
| Meeting  online  Meeting | Colloquium:Chemoinformaticsand ArtificialIntelligence  LatinXChem | Onlinemeetingforthegeneralpublicwithanemphasison undergraduateandgraduatestudents.  AvirtualforumthroughwhichthecommunityofLatin Americanchemistslocatedanywhereintheworldcanshare anddiscusstheirresearchresults.Itincludesasection  dedicatedtocomputationalchemistry.  TheaimwastointroduceresearcherstoChemoinformatics, especiallyinstructure-basedandligand-baseddrugdesign. Topicscovered:theuseofprotein,ligandanddrugdatabases, proteinmodeling,moleculardockingandvirtualscreeningfor drugdiscoveryapplications;managingandanalyzingvirtual chemicallibraries;machinelearningalgorithmsindrug  discovery.  Theobjectivewastointroducetheconceptsof  Chemoinformaticsanditsapplicationsintheacademic environment.Inaddition,shownationalresearcherstopics relatedtocheminformatics,datamining,database  representationsin2Dand3D.  Thisconferenceprovidedinformationonchemotherapyagainst LeishmaniawherethedevelopmentofCYP51inhibitordrugsis soughtasanewpurpose,soingeneraltermstheconference focusedontheuseofchemoinformaticsasacomplementto virtualdetectionapproaches(VS)basedonligandsand  structures.  Itincludedapplicationsofcomputationalchemistryto simulationforcarbonnanotubereinforcedcomposites. | [https://www.difacquim.com/english/events/ 2022-colloquium/[48]](https://www.difacquim.com/english/events/2022-colloquium/)  [https://www.latinxchem.org/lxchemcomp](https://www.difacquim.com/english/events/2022-colloquium/) |
| Workshop | CABANAWorkshop:  ChemoinformaticsinDrugDiscovery | https://www.ebi.ac.uk/training/events/  cabana-workshop-chemoinformatics-drug-discovery/ |
| Workshop | ComputationalChemistryinDrug DiscoveryandDevelopment.From AcademiatoIndustry.  Computer-AidedDrugDesign | InformationintheofficeofSponsorSENACYT. |
| Conference | IntegratedChemoinformatics  approachestovirtualscreeninginthe searchofnewleadcompounds  againstLeishmania | [https://www.researchgate.net/publication/](https://www.researchgate.net/publication/235412450_Integrated_chemoinformatics_approaches_to_virtual_screening_in_the_search_of_new_lead_compounds_against_Leishmania)  [235412450\_Integrated\_chemoinformatics\_](https://www.researchgate.net/publication/235412450_Integrated_chemoinformatics_approaches_to_virtual_screening_in_the_search_of_new_lead_compounds_against_Leishmania)  [approaches\_to\_virtual\_screening\_in\_the\_search\_of\_ new\_lead\_compounds\_against\_Leishmania](https://www.researchgate.net/publication/235412450_Integrated_chemoinformatics_approaches_to_virtual_screening_in_the_search_of_new_lead_compounds_against_Leishmania) |
| Conference | Informaticsmethodtobridgegap betweenexperimentalresultsand simulationforcarbonnanotube reinforcedcomposites  LatinAmericanChemistryand Catalysis.  WomeninBioinformatics&Data Science-LatinAmerica. | [https://repositorioslatinoamericanos.uchile.cl/ handle/2250/926817](https://repositorioslatinoamericanos.uchile.cl/handle/2250/926817) |
| Conference | Itincludedapplicationsofcomputationalchemistryto catalysis.  RecurrentmeetingbringingtogetherLatinAmerican  researcherswhoworkintheareasofsystemsbiology,omics technologies,artificialintelligence,machinelearning,data science,datamining,andhigh-performancecomputingwith applicationsinbiology.  Onlineeventdiscussingdevelopmentsandapplicationsand chemoinformatics,includingchemicalspace,dataanalysis, naturalproducts,virtualscreening,toxicityprediction,and drugdiscoveryforneglecteddiseases.Itincludedahands-on sessiontousetheChEMBL.Editorsofchemoinformatics journalsdiscussedpointsrelatedtoscientificpublishing.  Onlineeventtopublicizetheusesandpracticalapplicationsof informationandcommunicationstechnologiesduringthe pandemiccausedbytheCovid-19virus,whichincludedatalk ontheroleofChemoinformaticsandComputationalBiology duringthepandemic.  Websitewithtutorials,scriptlibraryandFAQfor pharmacoinformaticsanddrugdesign.  CourseopentoapplicantsfrominstitutesbasedinLatin AmericaandtheCaribbean.Theprogramwasaimedat researcherswithabackgroundknowledgeofanydiscipline relatedtodrugdiscovery,includingbiology,chemistry, pharmacology,computationalchemistryandinformatics. | [https:](https://www.youtube.com/watch?v=wlRbuIf2Hbw)  [//www.youtube.com/watch?v=wlRbuIf2Hbw https://wbds.la/[49]](https://www.youtube.com/watch?v=wlRbuIf2Hbw) |
| Conference |
| Conference | Schoolofchemoinformatics | [https:](https://www.youtube.com/@SchoolChemInfLA/videos)  [//www.youtube.com/@SchoolChemInfLA/videos](https://www.youtube.com/@SchoolChemInfLA/videos) |
| Conference | DigitalTechnologies,Innovationand HealthintheCovid-19Pandemic| PROSIC-UCR | [https://www.youtube.com/watch?v= T49FvEk6aCg™t=2196s](https://www.youtube.com/watch?v=T49FvEk6aCg&t=2196s) |
| Educational/ tutorial  Educational/ Course | RamirezLabGithub | <https://github.com/ramirezlab/WIKI> |
| PracticalAspectsofDrugDiscovery: AttheInterfaceofBiology,Chemistry andPharmacology-LatinAmerica andtheCaribbean | [https://coursesandconferences.](https://coursesandconferences.wellcomeconnectingscience.org/event/practical-aspects-of-drug-discovery-at-the-interface-of-biology-chemistry-and-pharmacology-latin-america-and-the-caribbean-20221113/)  [wellcomeconnectingscience.org/event/](https://coursesandconferences.wellcomeconnectingscience.org/event/practical-aspects-of-drug-discovery-at-the-interface-of-biology-chemistry-and-pharmacology-latin-america-and-the-caribbean-20221113/)  [practical-aspects-of-drug-discovery-at-the-](https://coursesandconferences.wellcomeconnectingscience.org/event/practical-aspects-of-drug-discovery-at-the-interface-of-biology-chemistry-and-pharmacology-latin-america-and-the-caribbean-20221113/) [interface-of-biology-chemistry-and-](https://coursesandconferences.wellcomeconnectingscience.org/event/practical-aspects-of-drug-discovery-at-the-interface-of-biology-chemistry-and-pharmacology-latin-america-and-the-caribbean-20221113/) [pharmacology-latin-america-and-the-caribbean-](https://coursesandconferences.wellcomeconnectingscience.org/event/practical-aspects-of-drug-discovery-at-the-interface-of-biology-chemistry-and-pharmacology-latin-america-and-the-caribbean-20221113/)[20221113/](https://difacquim.gitbook.io/quimioinformatica/)  [https://difacquim.gitbook.io/](https://difacquim.gitbook.io/quimioinformatica/)  [quimioinformatica/](https://difacquim.gitbook.io/quimioinformatica/) |
| Educational | OnlineManualofChemoinformatics (inSpanish). | Thismanualcontributestotheacademicformationofthe students,strengtheninginstudentstheunderstandingofbasic conceptsofChemoinformaticsandseekstoallowthemtobe abletohandleandinterpretcomputationaltechniques  associatedwiththisscientificdiscipline,frequentlyusedinthe discovery,designanddevelopmentofbioactivecompounds.  The13CNMRDatabaseofNaturalProductsisaweb  applicationthatprovidesdifferentsearchtoolstoknowthe structureofacompound.Inaddition,itallowstheuseof chemoinformaticsapplicationsbasedinJSME-X-2020byPeter ErtlandBrunoBeinfait,(Novartis),inthedesignand  developmentofundergraduatestudentthesesinPharmacy.  Meetingofresearchersandstudentsinterestedinacquiring knowledgeinthefollowingtopics:proteinstructureprediction, machinelearningmethods,molecularmodelingappliedto drugdevelopment,receptor-liganddockingmethodologies, moleculardynamics,montecarlo,quantumelectronicstructure calculations,hybridquantum/classicalcalculations,and  bioinformatics. |
| Educational  andResearch | NAPROC13 | <https://c13.materia-medica.net/> |
| Meeting  online | 10thSchoolofmolecularmodelingin biologicalsystems | <http://www.emmsb.lncc.br/> |

11

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*  *ArtificialIntelligenceintheLifeSciences3(2023)100077*

**Table5**

RepresentativestrategicactionstostrengthenchemoinformaticsinLatinAmerica.

|  |  |  |
| --- | --- | --- |
| Action | Timeframea | Expectedcontributionorimpact |
| Research:enhanceresearchcollaborationswithinLatin  AmericaandbetweenLatinAmericaandothercountries. Research:continueworkingontopicsrelevanttogeographical regionsandglobalissues.Foreseeupcomingchallenges,  emergingneeds,areasofopportunity,andtrends.  Research:Partnerwithindustry-drivenfoundationsand researchinstitutes(e.g.,trescantosOpenLabfoundation).  Research:Partnerwithsoftwarecompanies (e.g.,Pharmacelera).  Education:Promotetheuniversityresearchseedbedsin undergraduateprogramsatLatinAmericanuniversities. | Medium/long | Advancesignificantlyinresearchprojects(basicandappliedresearch),joining effortswithmulti-sites;increasethediversityofpointsofview.  Workingontopofregionalandglobalissuesandchallengeswillcontributeto maintainingLatinAmericancountriesontop. |
| Short/medium |
| Medium/long | Getinvolvedinbasicresearchprojectsthatcanbetranslatedintopractical applications.  Collaboratinginthedevelopmentofchemoinformaticstoolssuchasmolecular dockingprograms  Thiswillenhancetheinterestinresearchinundergraduatestudents,allowing themtogenerateskillsintheuseofcomputationaltoolsinthefacultiesof computerscience,biology,chemistry,andpharmacy.Thiswouldfostertheir interestinfurtheringstudiesinbasicchemoinformaticsanditsapplicationsin academia,industry,andemergingcompaniesinAI.  Itwillhelptotrainhighlyeducatedspecialistsinchemoinformaticsthatwill jointheindustry,academia,andstart-upcompanies.  Scientificconferenceshelpincreasetheinteractionsamongstudentsand researchscientistsintheareaandtheworld;willcontributetobeinguptodate withrecentdevelopmentsandidentifyingfutureneeds.  Expandfundingsourcesbeyondtraditionalgovernmentsupportwillhelpto increasetheresearchfundingitselfandwillpromoteworkonappliedprojects.  Studentsparticipatinginresearchprojectswillgainexperienceworkingon industry-drivenprojectsandhelpthemtoworkonindustrypositions.  *Idem*(previousentry) |
| Short/medium |
| Short/medium |
| Education:Establishandsustainpostgraduateprograms(e.g., Masters’s)inLatinAmericanuniversities.  Education:continueandincreasethenumberofquality conferences,seminars,andschools.Thesecouldbein-person, hybrid,orvirtualevents.  Funding:attractattention,collaboration,andpartnershipswith industries(pharmaceutical,food,materials,etc.)tofund  researchprogramswithtranslationalapplications. | Long |
| Short/medium |
| Medium/long |
| Funding:continueandincreasetheconnectionswith foundations,not-for-profitinstitutions,publishers,and industriesintheregionandworldwidethathelpsupport educationalandresearchprograms.  General:CreateaLatinAmericanchemoinformaticssociety thatguidesandoverseesthedevelopment.  Disseminationandcollaborations:continuetoenhancethe presenceandparticipationofLatinAmericanScientistsin internationalconferencesorganizedinotherpartsoftheworld. | Medium |
| Long | Helptosustainthecollectiveeffortsintermsofresearch,education,and funding.  ActiveparticipationofLatinAmericanstudentsandscientistswillfavorthe generationand/orconsolidationofcollaborativenetworks;willalsohelpthe awarenessofthescientificcommunityworldwideoftheresearchdoneinLatin America.  StrengthentheconnectionandnetworksbetweenLatinAmericawithother countries,be“ambassadors” thatsharetheneedsandprogressintheLatin Americanregion. |
| Short/medium |
| Disseminationandcollaborations:Continueandenhancethe participationofLatinAmericanscientistsinspecialized  internationalcommitteesandeditorialboardsofpeer-reviewed journals.  Chemoinformaticscommunity-wideblindchallenges. | Medium |
| Long | Withtheprogressinthegenerationofdatabasesofnaturalproducts,biological activitiesandphysicochemicalpropertiesofcompoundsintheregion,itis intendedtodemocratizereliabledatasothatthescientificcommunityaround theworldcanvalidatecomputationaltools,proposenewpredictivemodelsthat canbecriticallyevaluatedthroughblindchallenges. |
| aApproximatetimeframetostarttheimplementationanddevelopmentofthetask.Inallcases,itisexpectedthattheactionsustainsforseveralyears. Short-term:1year;medium-term:1–5years;long-term:5–10years. | | |

interrelationships.Amanualofthedisciplineisevenpresented,with whichanintegrationofagreaternumberofscientistswhoapproach thisdisciplineisgenerated.Conferencesandwebinarsareveryillustra-tiveoftheapplicabilitythatisbeinggiventochemoinformatics,allow-inginteractionswithexperts,whocanguidetowardstheareasoffocus andinterestoftheviewer.Sincemanywebinarswererecordedandare freelyavailable,theyarealsousefulresourcesforteachingandtrain-ing.Atthesametime,theyareexcellentreferencestowitnessandkeep trackoftheevolutionofchemoinformaticsduringthenextfewyears. Althoughafewexamplesarelisted,theconferences,workshopsandre-sourcesinTable4pointtotheeffortsofseveralcountriestodevelop chemoinformaticsinLatinAmerica.

DuetotherecentCOVID-19pandemic,theonlineandhybridfor-matscouldrepresentanadvantagetoincreasethenumberofnational, regional,and/orinternationalconferences.Ofcourse,suchonlineorhy-bridconferenceswouldbeaddedtothefullin-personmeetingsthatalso providekeyadvantagesthroughthedirectinteractionbetweenscientists andstudents.

**5.Futuredirections**

Inthissection,wesummarizetheauthor’sperspectiveontheex-pectedprogressandevolutionofchemoinformaticsinLatinAmerica duringthenextfewyears.BasedonthediscussionoftheprecedingSec-

12

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*

versityresearchhotbedswithgroupsarealreadywellestablished.In anycase,face-to-face(orhybrid)meetingsthatarealsobeneficialcan beconductedandfinancedthroughcollaborationswithinLatinAmer-ica.Theseinteractionsleadustothegenerationofconglomeratesthat candevelopmulti-centerprojectsintheLatinAmericanregion.

Thiswillfavortheinterestsoffinancingagencies(publicorprivate) infinancingcollaborativeprojects.Theabove,togetherwiththeadvan-tagesofonlineandliveconferencesandwebinars,makeitincreasingly easiertoholdmeetingsorconferencesthathavetheadvantageofsaving timeandreducingthecostoforganizationandattendance.Itisalsoben-eficialtohavethecontinuedconvenienceofbeinginvolvedineditorial boardsofpeer-reviewedpublications.Currently,severalLatinAmeri-cansaretakinganactivepartintheeditorialboardsofjournalssuchas theJournalofChemoinformatics,JournalofChemicalInformationand Modeling,eLife,tonameafewexamples.Also,acurrenttrendofsome scientificpublishersistoincludeearlycareerresearchersinthereview board,e.g.,ProteinScience,whichopensopportunitiesforscientistsin LatinAmericatotakepartineditorialboardswhereourrepresentation isunderrepresented.

InseveraloftheactionpointsproposedinTable5,itiscrucialto expeditethepaperworkneededtoformalizeconfidentialagreements withindustry.Theoverabundanceofbureaucracyfrequentlysapsin-dustryandotherinstitutions’motivationtocollaboratewithacademic organizationsonscientificachievements.Ofcourse,thesepointsre-quireagreatereffortthatincludestheparticipationofhighmanage-mentandgovernments.Inall,research,education,andfundingare amongthemainpillarsthatwillhelptosustainandimproveChemoin-formaticsandrelateddisciplinesinLatinAmerica.Theresearchin-cludesdisseminationthroughpeer-reviewedpublications,participation ininternationalscientificconferences,andcreationandstrengthen-ingofcollaborationnetworkswithinLatinAmericaandwithother countries.

**6.Summaryconclusions**

Thisisacollectiveefforttodiscussthestatus,progress,andchal-lengesofchemoinformaticsinLatinAmericabyexpertsworkingonthis geographicalregion.AlthoughthereareseveralotherLatinAmerican researchgroupsworkingonchemoinformaticsthatarenotincludedin thismanuscript,weconsiderthatthisisthefirstefforttoteamupand discussthesubjectthatislooselybasedonthe“fourW´s”:whenthe publicationsandapplicationshavebeendisclosed(forthisworkwefo-cusedontheperiod2010–2022);who(e.g.,LatinAmericancountries workingonchemoinformaticsalongwithnetworkcollaborationswithin thecontinentandtherestoftheworld);whatapplicationsofchemoin-formaticsarebeingpublished;andwhere(weanalyzedthejournals, onlinedatabases,scientificconferences,andforumswheretheresearch isbeingdisseminated).Weconcludedthatlike“bioinformatics,” there isaneedtoestablishauniquetermforawell-definedandindepen-dentdiscipline.Itisalsoconcludedthat,overall,LatinAmericaisthe fourthgeographicalregionthatcontributestochemoinformaticsworld-wideafterEurope,Asia,andNorthAmerica.Byfar,BrazilistheLatin Americancountrywiththestrongestcontributions,followedbyMex-ico.Inthelasttwelveyears,themostcommoncheminformaticsap-plicationswithotherdisciplineshavebeenmolecularmodelingand drugdiscoverythat,ingeneralarecommonapplicationsofchemoinfor-maticinothergeographicalregions.Also,QSARandchemicalspaceare commontopicspublishedalongwithchemoinformatics.Collectively, MolecularInformaticsandJournalofChemoinformatics,followedby theJournalofChemicalInformationandModeling,arethemostfre-quentjournalswhereLatinAmericangroupspublishtheirpapers.The sameconclusioncanbedrawnfromthescientificjournalswherere-searchersfromothercountriespublishtheirchemoinformaticresearch papers.

Wealsoconcludethatthereisastrongcollaborationnetworkwithin LatinAmerica,thestrongestbeingbetweenBrazilandMexicothisfar.

13

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*

**Funding**

WethanktheDirecciónGeneraldeCómputoydeTecnologíasde InformaciónyComunicación(DGTIC),UNAM,forthecomputational resourcestouseMiztlisupercomputeratUNAMunderprojectLANCAD-UNAM-DGTIC-335;andtheinnovationspaceUNAM-HUAWEIthecom-putationalresourcestousetheirsupercomputerunderprojectNo.7“De-sarrolloyaplicacióndealgoritmosdeinteligenciaartificialparaeldis-eñodefármacosaplicablesaltratamientodediabetesmellitusycáncer”. WealsothankthefundingofDGAPA,UNAM,ProgramadeApoyoa ProyectosdeInvestigacióneInnovaciónTecnológica(PAPIIT),grant No.IN201321.Thisresearchwasalsofundedby:FONDECYT(Fondo NacionaldeDesarrolloCientíficoyTecnológico—Chile),grantnum-ber1220656andANID(AgenciaNacionaldeInvestigaciónyDesar-rollo—Chile)grantnumberATE220043andFOVI210027;ANID– MILE-NIO– NCN19\_038;Vice-rectoryofResearchandPostgraduateStudies oftheUniversityofPanamaforUniversityResearchFundsCUFI-VIP-01–14–2019–05andSNIsponsor.WZthankstheViceChancellorfor ResearchoftheUniversityofCostaRicaforitssupportworkviathe researchprojects115-C2–126and908-C3–610.Brazilianresearchers wasfundedbyConselhoNacionaldeDesenvolvimentoCientíficoeTec-nológico(CNPq-303757/2019–1and311875/2022–0),CAPESfounda-tion,grantnumbers88887.684031/2022–00;UFMGintramuralfunds anditsPró-ReitoriadePós-Graduação(PRPG)intramuralfunds.

**Supplementarymaterials**

Supplementarymaterialassociatedwiththisarticlecanbefound,in theonlineversion,at[doi:10.1016/j.ailsci.2023.100077](https://doi.org/10.1016/j.ailsci.2023.100077).

**References**

[1]VarnekA,BaskinII.[Chemoinformaticsasathe](https://doi.org/10.1002/minf.201000100)oreticalchemistrydiscipline.MolInf 2011;30:20–32.doi[:10.1002/minf.201000100.](https://doi.org/10.1002/minf.201000100)

[[2]López-LópezE,Baj](https://doi.org/10.1021/acs.jcim.0c01301)[orathJ,Medina-FrancoJL](https://doi.org/10.1002/minf.201000100)[.Informaticsforchemistry,biology, andbiomedicalsciences.JChemInfModel2021;61:26–35.doi:10.1021/acs.jcim. 0c01301.](https://doi.org/10.1021/acs.jcim.0c01301)

[[3]BajorathJ.Understandingchemoinformatics:aunifyingapproach.DrugDiscovTo-](https://doi.org/10.1021/acs.jcim.0c01301) day2004;9:13–14.doi[:10.1016/s1359-6446(04)02916-2.](https://doi.org/10.1016/s1359-6446(04)02916-2)

[4]GasteigerJ,FunatsuK[.Chemoinformatics– Animporta](https://doi.org/10.1016/s1359-6446(04)02916-2)ntscientificdiscipline.J ComputChem2006;5(2):53–8.doi[:10.2477/jccj.5.53.](https://doi.org/10.2477/jccj.5.53)

[5]GasteigerJ.Chemistry[intime](https://doi.org/10.1002/cphc.202000518)[sofartificia](https://doi.org/10.2477/jccj.5.53)l[i](https://doi.org/10.2477/jccj.5.53)ntelligence.Chemphyschem 2020;21(20):2233–42.doi[:10.1002/cphc.202000518.](https://doi.org/10.1002/cphc.202000518)

[[6]Medina-FrancoJL,Saldív](https://doi.org/10.3390/biom10111566)[ar-GonzálezFI.Cheminfor](https://doi.org/10.1002/cphc.202000518)[maticstocharacterizephar-macologicallyactivenaturalproducts.Biomolecules2020:1566.doi:10.3390/ biom10111566.](https://doi.org/10.3390/biom10111566)

[[7]GasteigerJ.Chemoinformatic](https://doi.org/10.3390/biom10111566)[s:achievementsandchalleng](https://doi.org/10.3390/molecules21020151)[es,apersonalview.](https://doi.org/10.3390/biom10111566)  Molecules2016;21(2):151.doi[:10.3390/molecules21020151.](https://doi.org/10.3390/molecules21020151)

[8]Medina-FrancoJL,etal.Bridg[inginformaticsandmedicina](https://doi.org/10.3390/molecules21020151)linorganicchemistry: towardadatabaseofmet[allodrugsandmetallodrugcan](https://doi.org/10.1016/j.drudis.2022.02.021)didates.DrugDiscovToday 2022;27(5):1420–30.doi[:10.1016/j.drudis.2022.02.021.](https://doi.org/10.1016/j.drudis.2022.02.021)

[9]KirchmairJ.Molecular[informaticsinnaturalpro](https://doi.org/10.1016/j.drudis.2022.02.021)ductsresearch.MolInf 2020;39(11):e2000206.doi[:10.1002/minf.202000206.](https://doi.org/10.1002/minf.202000206)

[10]NavejaJJ,Oviedo-Osornio[CI,Trujillo-MineroNN,M](https://doi.org/10.1002/minf.202000206)edina-FrancoJL.Chemoin-formatics:aperspective[fromanacademicsettingin](https://doi.org/10.1007/s11030-017-9802-3)LatinAmerica.MolDiversity 2017;22(1):247–58.doi[:10.1007/s11030-017-9802-3.](https://doi.org/10.1007/s11030-017-9802-3)

[[11]Perianes-RodriguezA,e](https://doi.org/10.1016/j.joi.2016.10.006)[tal.Constructingbibliometric](https://doi.org/10.1007/s11030-017-9802-3)[networks:acomparisonbe-tweenfullandfractionalcounting.JInformetr2016;10(4):1178–95.doi:10.1016/ j.joi.2016.10.006.](https://doi.org/10.1016/j.joi.2016.10.006)

[[12]González-MedinaM,JesúsNavejaJ,Sánchez-CruzN,Medina-FrancoJL.Open](https://doi.org/10.1016/j.joi.2016.10.006) chemoinformaticresourcestoexplorethestructu[re,propertiesandchem](https://doi.org/10.1039/C7RA11831G)icalspace ofmolecules.RSCAdv2017;7(85):54153–63.doi[:10.1039/C7RA11831G.](https://doi.org/10.1039/C7RA11831G)

[[13]CoxPB,GuptaR.Contemporarycomputationala](https://doi.org/10.1021/acsmedchemlett.1c00662)[pplicationsandtoolsin](https://doi.org/10.1039/C7RA11831G)[drugdis-covery.ACSMedChemLett2022;13(7):1016–29.doi:10.1021/acsmedchemlett. 1c00662.](https://doi.org/10.1021/acsmedchemlett.1c00662)

[[14]VilloutreixBO,etal.Onehundredthousandmouseclicksdowntheroad:selected](https://doi.org/10.1021/acsmedchemlett.1c00662) onlineresourcessupportingdrugd[iscoverycollectedoveradeca](https://doi.org/10.1016/j.drudis.2013.06.013)de.DrugDiscov Today2013;18(21–22):1081–9.doi[:10.1016/j.drudis.2013.06.013.](https://doi.org/10.1016/j.drudis.2013.06.013)

[15]Gómez-GarcíaA,Medina-FrancoJL.[ProgressandimpactofLatinA](https://doi.org/10.1016/j.drudis.2013.06.013)[mericannatu](https://doi.org/10.3390/biom12091202)ral productdatabases.Biomolecules2022;12(9):1202.doi[:10.3390/biom12091202.](https://doi.org/10.3390/biom12091202) [16]LMedina-FrancoJ.TowardsaunifiedLatinAmerican[naturalproductsdataba](https://doi.org/10.3390/biom12091202)se: [lANaPD.FutureSciOA2020;6(8):FSO468.doi](https://samplchallenges.github.io/)[:10.2144/fsoa-2020-0068.](https://doi.org/10.2144/fsoa-2020-0068)

[[17]Mobley,D.L.TheSAMPLchallenges.SAMPL.2](https://samplchallenges.github.io/)[023https://samplchalleng](https://doi.org/10.2144/fsoa-2020-0068)[es.github.](https://samplchallenges.github.io/)  [io/(accessed2023-04-30).](https://samplchallenges.github.io/)

[[18]GeballeMT,GuthrieJP.TheSAMPL3blindpredictionchallenge:transfer](https://samplchallenges.github.io/) [energyoverview.JComput-AidedMolDes2012;26(5):489–96.doi:10.1007/ s10822-012-9568-8.](https://doi.org/10.1007/s10822-012-9568-8)

*ArtificialIntelligenceintheLifeSciences3(2023)100077*

[19]MobleyDL,WymerKL,LimNM,GuthrieJP.Blindpredictionofsolvationfreeen-erg[iesfromtheSAMPL4challeng](https://doi.org/10.1007/s10822-014-9718-2)e.JComput-AidedMolDes2014;28(3):135–50. doi[:10.1007/s10822-014-9718-2.](https://doi.org/10.1007/s10822-014-9718-2)

[20]Işık[M,RustenburgAS,RizziA,](https://doi.org/10.1007/s10822-014-9718-2)GunnerMR,MobleyDL,ChoderaJD.Overview [oftheSAMPL6pKchallenge:evaluatingsmallmoleculemicroscopicandmacro-scopicpKpredictions.JComput-AidedMolDes2021;35(2):131–66.doi:10.1007/ s10822-020-00362-6.](https://doi.org/10.1007/s10822-020-00362-6)

[21][BergazinTD,TielkerN,ZhangY,MaoJ,GunnerMR,FranciscoK,BallatoreC,](https://doi.org/10.1007/s10822-020-00362-6) [KastSM,MobleyDL.EvaluationoflogP,pK,andlogDpredictionsfromthe SAMPL7blindchallenge.JComput-AidedMolDes2021;35(7):771–802.doi:10. 1007/s10822-021-00397-3.](https://doi.org/10.1007/s10822-021-00397-3)

[22][IşıkM,BergazinTD,FoxT,RizziA,ChoderaJD,MobleyDL.Assessingthe](https://doi.org/10.1007/s10822-021-00397-3) [accuracyofoctanol-waterpartitioncoefficie](https://doi.org/10.1007/s10822-020-00295-0)nt[predictionsintheSAMPL6part IIlogPChallenge.JComput-AidedMolDes2020;34(4):335–70.doi:10.1007/ s10822-020-00295-0.](https://doi.org/10.1007/s10822-020-00295-0)

[23][O’BoyleNM,BanckM,JamesCA,MorleyC,VandermeerschT,HutchisonGR.](https://doi.org/10.1007/s10822-020-00295-0)

[Openbabel:anopenchemicaltoolbox.JCheminf2011;3:33.doi:10.1186/ 1758-2946-3-33.](https://doi.org/10.1186/1758-2946-3-33)

[24][BarreiroEJ,RodriguesCR,AlbuquerqueMG,Sant’AnnaCMR,AlencastroRB.](https://doi.org/10.1186/1758-2946-3-33)

[modelagemmolecular:umaferramentaparaoplanejamentoracionaldefár-macosemquímicamedicinal.QuímNova1997;20(3):300–10.doi:10.1590/ S0100-40421997000300011.](https://doi.org/10.1590/S0100-40421997000300011)

[25][MontanariCA.Aquímicamedicinalnapróximadécada.Quím.Nova](https://doi.org/10.1590/S0100-40421997000300011)  2000;23(1):134–7.doi[:10.1590/S0100-40422000000100024.](https://doi.org/10.1590/S0100-40422000000100024)

[26]GóesACS,OliveiraBV[X.Projetogenomahumano:umretrato](https://doi.org/10.1590/S0100-40422000000100024)daconstruçãodocon-hecimentocient[ífic](https://doi.org/10.1590/1516-73132014000300004)o[sobaóticadarevistaciênciahoj](https://doi.org/10.1590/1516-73132014000300004)e.Ciênceduc2014;20(3):561–77(Bauru).doi[:10.1590/1516-73132014000300004.](https://doi.org/10.1590/1516-73132014000300004)

[27][FerreiraLG,D](https://doi.org/10.3390/molecules200713384)[osSantosRN,OlivaG,Andricopulo](https://doi.org/10.1590/1516-73132014000300004)[AD.Moleculardockingand structure-baseddrugdesignstrategies.Molecules2015;20(7):13384–421.doi:10. 3390/molecules200713384.](https://doi.org/10.3390/molecules200713384)

[28][Tavare,A.;Horta,C.;Kümmerlec,A.E.;Guido,R.Aevoluçãodaquímicamedicinal](https://doi.org/10.3390/molecules200713384) no[Brasil:avançosnos40anosdasoci](https://doi.org/10.21577/0100-4042.20170075)edadebrasileiradequímica.2017;40:694–700 doi[:10.21577/0100-4042.20170075.](https://doi.org/10.21577/0100-4042.20170075)

[29]Am[aralAT,MontanariCA.Química](https://doi.org/10.21577/0100-4042.20170075)medic[inal:25anosdeplanejamentoracional](https://doi.org/10.1590/S0100-40422002000800008)de fármacos.QuímNova2002;25:39–44.doi[:10.1590/S0100-40422002000800008.](https://doi.org/10.1590/S0100-40422002000800008) [30][Sant’AnnaCMR.Glossáriodetermosus](https://doi.org/10.1590/S0100-40422002000300027)[adosnoplanejamentodefármacos(r](https://doi.org/10.1590/S0100-40422002000800008)[e-](https://doi.org/10.1590/S0100-40422002000300027) [comendaçõesdaIUPACpara1997).QuímNova2002;25(3):505–12.doi:10.1590/](https://doi.org/10.1590/S0100-40422002000300027)  [S0100-40422002000300027.](https://doi.org/10.1590/S0100-40422002000300027)

[31][VerliH,BarreiroEJ.Umparadigmadaquímicamedicinal:aflexibili](https://doi.org/10.1590/S0100-40422002000300027)-[dadedosligantesereceptores.QuímNova2005;28(1):95–102.doi:10.1590/ S0100-40422005000100018.](https://doi.org/10.1590/S0100-40422005000100018)

[32][LimaLM.Químicamedicinalmoderna:desafio](https://doi.org/10.1590/S0100-40422005000100018)s[econtribuiçãobrasileira.Quím](https://doi.org/10.1590/S0100-40422005000100018)  [Nova2007;30(6):1456–68.doi](https://www.chemistryworld.com/news/encoding-creativity-in-drug-discovery/4016866.article)[:10.1590/S0100-40422007000600015.](https://doi.org/10.1590/S0100-40422007000600015)

[33][King,A.Encodingcreativityind](https://www.chemistryworld.com/news/encoding-creativity-in-drug-discovery/4016866.article)[rugdiscovery.ChemistryWorld.2023](https://doi.org/10.1590/S0100-40422007000600015)[https://www.](https://www.chemistryworld.com/news/encoding-creativity-in-drug-discovery/4016866.article)

[chemistryworld.com/news/encoding-creativity-in-drug-discovery/4016866.article (accessed2023-04-13).](https://www.chemistryworld.com/news/encoding-creativity-in-drug-discovery/4016866.article)

[34]GoriDNP,AlbercaLN,RodriguezS,AliceJI,LlanosMA,BelleraCL,TaleviA.LIDeB tools:aLatinAmericanresourceoffreelyav[ailable,open-sourcecheminfo](https://doi.org/10.1016/j.ailsci.2022.100049)rmatics apps.ArtifInte[llLifeSci2022;2:100049.doi](https://naturar.quimica.unlp.edu.ar/es/)[:10.1016/j.ailsci.2022.100049.](https://doi.org/10.1016/j.ailsci.2022.100049)

[35]Website.2023[https://naturar.quimica.unlp.](https://naturar.quimica.unlp.edu.ar/es/)[edu.ar/es/.](https://doi.org/10.1016/j.ailsci.2022.100049)

[36]ValliM,dosS[antosRN,FigueiraLD,NakajimaCH,C](https://naturar.quimica.unlp.edu.ar/es/)astro-GamboaI,Andricop-uloAD,BolzaniVS.Developmentofanaturalpro[ductsdatabasefromt](https://doi.org/10.1021/np3006875)hebiodiver-sityofBrazil.JNatProd2013;76(3):439–44.doi[:10.1021/np3006875.](https://doi.org/10.1021/np3006875)

[37]PilonAC,ValliM,DamettoAC,PintoMEF,Freir[eRT,Castro-Gamboa](https://doi.org/10.1021/np3006875)I,Andricop-[uloAD,BolzaniVS.NuBBE:anupdateddatabasetouncoverchemicalandbiolog-icalinformationfromBrazilianbiodiversity.SciRep2017;7(1):7215.doi:10.1038/ s41598-017-07451-x.](https://doi.org/10.1038/s41598-017-07451-x)

[38][ScottiMT,Herrera-AcevedoC,OliveiraTB,CostaRPO,SantosSYKO,RodriguesRP,](https://doi.org/10.1038/s41598-017-07451-x) [ScottiL,Da-CostaFB.SistematX,anonlineweb-basedcheminformaticstoolfor datamanagementofsecondarymetabolites.Molecules2018;23(1).doi:10.3390/ molecules23010103.](https://doi.org/10.3390/molecules23010103)

[39][CostaRPO,LucenaLF,SilvaLMA,ZocoloGJ,Herrera-AcevedoC,ScottiL,Da-](https://doi.org/10.3390/molecules23010103)[CostaFB,IonovN,PoroikovV,MuratovEN,ScottiMT.TheSistematXwebportalof naturalproducts:anupdate.JChemInfModel2021;61(6):2516–22.doi:10.1021/](https://doi.org/10.1021/acs.jcim.1c00083) [acs.jcim.1c00083.](http://zinc12.docking.org/catalogs/uefsnp)

[40][ZINCcatalogUEFSnaturalproducts.2023http://zinc12.docking.org/catalogs/](http://zinc12.docking.org/catalogs/uefsnp)  [uefsnp(accessed2023-03-15.](http://zinc12.docking.org/catalogs/uefsnp)

[41][Zamora,W.J.;Pinheiro,S.;Acuña,D.NAPRORE-CR(NAturalPROductsREpository](http://zinc12.docking.org/catalogs/uefsnp) [-CostaRica),2023.doi](https://uniiquim.iquimica.unam.mx/)[:10.5281/zenodo.7858102.](https://doi.org/10.5281/zenodo.7858102)

[42][Unidaddeinformática](https://uniiquim.iquimica.unam.mx/)[delInstitutodeQuímica.](https://doi.org/10.5281/zenodo.7858102)[2023https://uniiquim.iquimica.](https://uniiquim.iquimica.unam.mx/)  [unam.mx/(accessed2023-03-15).](https://uniiquim.iquimica.unam.mx/)

[43][Pilón-JiménezBA,Saldívar-GonzálezFI,Díaz-EufracioBI,Medina-FrancoJL.BIO-](https://uniiquim.iquimica.unam.mx/)FACQUIM:a[Mexicancompoundda](https://doi.org/10.3390/biom9010031)tabaseofnaturalproducts.Biomolecules 2019;9(1).doi[:10.3390/biom9010031.](https://doi.org/10.3390/biom9010031)

[44]Sánchez-Cruz[N,Pilón-JiménezBA,M](https://doi.org/10.3390/biom9010031)edina-FrancoJL.Functionalgroupanddi-versityanal[ysisofBIOFACQUIM:aMexican](https://doi.org/10.12688/f1000research.21540.2)naturalproductdatabase.F1000Res 2019;8.doi[:10.12688/f1000research.21540.2.](https://doi.org/10.12688/f1000research.21540.2)

[45]OlmedoDA[,González-MedinaM,GuptaMP,](https://doi.org/10.12688/f1000research.21540.2)Medina-FrancoJL.Cheminformatic charact[erizationofnaturalproductsf](https://doi.org/10.1007/s11030-017-9781-4)romPanama.MolDiversity2017;21(4):779–89.doi[:10.1007/s11030-017-9781-4.](https://doi.org/10.1007/s11030-017-9781-4)

[46]Barazor[da-Ccahuana,H.L.,Ranilla,L](https://doi.org/10.1007/s11030-017-9781-4).G.,Candia-Puma,M.A.,Cárcamo-Rodriguez, E.G.,Centeno-Lopez,A.E.,Del-Carpio,G.D.,Medina-Franco,J.L.,Chávez-Fumagalli, M.A.PeruNPDB:thePe[ruviannaturalproductsdatabaseforinsilicodr](https://doi.org/10.1038/s41598-023-34729-0)ugscreening.

*[SciRep](http://refhub.elsevier.com/S2667-3185(23)00021-1/sbref0047)*[202313,7577.](http://refhub.elsevier.com/S2667-3185(23)00021-1/sbref0047)<https://doi.org/10.1038/s41598-023-34729-0>  
[47][SwamyMK,KumarA.](http://refhub.elsevier.com/S2667-3185(23)00021-1/sbref0047)[Phytochemicalgenomics:plantmetabolomicsa](https://doi.org/10.1038/s41598-023-34729-0)[ndmedicinal](http://refhub.elsevier.com/S2667-3185(23)00021-1/sbref0047)  [plantgenomics.SpringerNature;2023.](http://refhub.elsevier.com/S2667-3185(23)00021-1/sbref0047)

14

*J.Miranda-Salas,C.Peña-Varas,I.ValenzuelaMartínezetal.*

[48]BajorathJ,Chávez-HernándezAL,Duran-FrigolaM,Fernández-deGortariE, GasteigerJ,López-LópezE,MaggioraGM,Medina-FrancoJL,Méndez-LucioO, MestresJ,Miranda-QuintanaRA,OpreaTI,PlissonF,Prieto-MartínezFD, Rodríguez-PérezR,Rondón-VillarrealP,Saldívar-GonzalezFI,Sánchez-CruzN, [ValliM.Chemoinformaticsandartificia](https://doi.org/10.1186/s13321-022-00661-0)l[intelligencecolloquium:progressandchal-lengesindevelopingbioactivecompounds.JCheminf2022;14(1):1–12.doi:10. 1186/s13321-022-00661-0.](https://doi.org/10.1186/s13321-022-00661-0)

[[49]VelezRuedaAJ,etal.Womeninbioinformatics&datascience-LatinAmerica.](https://doi.org/10.1186/s13321-022-00661-0)  MethodsX2022;9:101907.doi[:10.1016/j.mex.2022.101907.](https://doi.org/10.1016/j.mex.2022.101907)

[50]BannanCC,BurleyKH,ChiuM,[ShirtsMR,GilsonMK,Mobley](https://doi.org/10.1016/j.mex.2022.101907)DL.Blindpredictionof cyclohexane-waterdistributioncoefficien[tsfromtheSAMPL5Challeng](https://doi.org/10.1007/s10822-016-9954-8)e.JComput-AidedMolDes2016;30(11):927–44.doi[:10.1007/s10822-016-9954-8.](https://doi.org/10.1007/s10822-016-9954-8)

[51]RiquelmeM,Vöhringer-MartinezE.SA[MPL6octanol-waterpartition](https://doi.org/10.1007/s10822-016-9954-8)coefficients fromalchemicalfreeenergycalculatio[nswithMBISAtomicCharges](https://doi.org/10.1007/s10822-020-00281-6).JComput-AidedMolDes2020;34(4):327–34.doi[:10.1007/s10822-020-00281-6.](https://doi.org/10.1007/s10822-020-00281-6)

*ArtificialIntelligenceintheLifeSciences3(2023)100077*

[52]ZamoraWJ,PinheiroS,GermanK,RàfolsC,CurutchetC,LuqueFJ.Prediction ofthen-octanol/waterpartitioncoefficientsintheSAMPL6blindchallengefrom MSTco[ntinuumsolvationcalculations](https://doi.org/10.1007/s10822-019-00262-4).JComput-AidedMolDes2020;34(4):443–51.doi[:10.1007/s10822-019-00262-4.](https://doi.org/10.1007/s10822-019-00262-4)

[53]Viayna[A,PinheiroS,CurutchetC,](https://doi.org/10.1007/s10822-019-00262-4)LuqueFJ,ZamoraWJ.Predictionofn-octanol/waterpartitioncoefficientsandacidityconstants(pK)intheSAMPL7blind challen[gewiththeIEFPCM-MSTmode](https://doi.org/10.1007/s10822-021-00394-6)l.JComput-AidedMolDes2021;35(7):803–11.doi[:10.1007/s10822-021-00394-6.](https://doi.org/10.1007/s10822-021-00394-6)

[54]Rodrig[uezSA,TranJV,SabatinoSJ,](https://doi.org/10.1007/s10822-021-00394-6)PaluchAS.Predictingoctanol/waterparti-[tioncoefficien](https://doi.org/10.1007/s10822-022-00474-1)ts[andpKafortheSAMPL7challengeusingtheSM12,SM8and SMDsolvationmodels.JComput-AidedMolDes2022;36(9):687–705.doi:10.1007/ s10822-022-00474-1.](https://doi.org/10.1007/s10822-022-00474-1)

[55][LopezK,PinheiroS,ZamoraWJ.Multiplelinearregressionmodelsforpredicting](https://doi.org/10.1007/s10822-022-00474-1) then‑octanol/waterpartitioncoefficien[tsintheSAMPL7blindchalleng](https://doi.org/10.1007/s10822-021-00409-2)e.JComput-AidedMolDes2021;35(8):923–31.doi[:10.1007/s10822-021-00409-2.](https://doi.org/10.1007/s10822-021-00409-2)

15