UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE ENGENHARIA DE ILHA SOLTEIRA
DEPARTAMENTO DE FÍSICA E QUÍMICA

RELATÓRIO DE PESQUISA DE PÓS-DOUTORADO

Engenharia de cátions do sítio A em perovskitas de haleto

Proponente: Fernando Brondani Minussi

Supervisor: Prof. Dr. Eudes Borges de Araújo

Ilha Solteira - SP

Março, 2026

Resumo

As perovskitas de haleto são uma classe de materiais com aplicações optoeletrônicas e eletroquímicas consolidadas, alcançando eficiências comparáveis às de materiais estabelecidos nos respectivos campos. Entretanto, este cenário de maturidade tecnológica foi, em grande parte, resultado do conhecimento empírico na área, sem que as relações composição-estrutura-propriedade tenham sido plenamente entendidas. Em composições de cátions mistos, por exemplo, geralmente se observam algumas relações diretas entre os tamanhos dos substituintes e as propriedades das perovskitas. Porém, em vários casos, estas relações não são observadas, implicando que outras características destes cátions desempenham um papel importante. Apesar de sua importância, esse entendimento não foi totalmente aprofundado. Para estudar esse problema, realizou-se a síntese e caracterização de mais de 65 composições de cátions mistos no sítio A baseadas em misturas entre metilamônio, acetamidínio, formamidínio, etilamônio, guanidínio e trimetilsulfônio. Os resultados experimentais foram agrupados para estudar a correlação composição-propriedades envolvendo características estruturais, optoeletrônicas e elétricas. Esses dados, em conjunto com da literatura, foram usados para formar um banco de dados com informação de fases de mais de 200 composições contendo esses e outros cátions no sítio A para previsão da temperatura de transição de fases em compostos do tipo $APbI_3$. Um estudo análogo foi iniciado envolvendo a previsão de energias de *band gap* em composições quaisquer de perovskitas de haleto. Em todos os casos, foram utilizadas abordagens estatísticas e/ou de *machine learning* para obtenção de relações qualitativas e quantitativas que permitem a engenharia de cátions mistos do sítio A para se atingir propriedades de interesse em perovskitas de haleto ainda não sintetizadas.

1. INTRODUÇÃO

Perovskitas de haleto são compostos com a composição geral ABX_3, onde A representa um cátion monovalente, B um cátion divalente e X um ânion haleto, construídos por um arranjo tridimensional de octaedros BX_6 de compartilhamento de vértices [1]. Em virtude da sua vastidão química, estrutural, sintética e de propriedades, essa classe de materiais encontra aplicações optoeletrônicas e eletroquímicas diversas, como em células solares, baterias, capacitores, LEDs, detectores de radiação, conversão termoelétrica, sensores de gases, produção de hidrogênio e dispositivos neuromórficos [2-20]. Com o intenso progresso inicial e o atual estabelecimento dos fundamentos desses materiais, com altas eficiências sendo obtidas em dispositivos de perovskita, como em células solares [21], esses materiais podem estar próximos da comercialização [22,23], pois seu principal gargalo, isto é, seus problemas de instabilidade, continuam sendo mais bem abordados [24-30]. Nesse cenário de maturidade tecnológica, grande parte dos desafios científicos atuais reside na consolidação do entendimento dos princípios por trás do conhecimento empírico sobre esses materiais. Essas mesmas preocupações foram apontadas recentemente em relação à foto-estabilidade de perovskitas de haleto [31].

A formulação de soluções sólidas para cátions mistos no sítio A é uma das estratégias relevantes para obter o conjunto desejado de características em perovskitas de haleto [32-34]. De fato, a maioria dos dispositivos de alta eficiência fabricados com perovskitas de haleto usam composições de cátions mistos [35,36]. Nessas composições mistas de sítio A, uma infinidade de cátions pode ser usada, incluindo inorgânicos como potássio (K^+), rubídio (Rb^+) e césio (Cs^+), e orgânicos como metilamônio (MA^+), formamidínio (FA^+), imidazólio (IM^+), hidrazínio (Hy), metil-hidrazínio (MHy^+), aziridínio (AZi^+), azetidínio (AZe^+), dimetilamônio (DMA^+), etilamônio (EA^+), acetamidínio (AC^+) e guanidínio (GA^+) [37-42]. As regras gerais para a formação de perovskitas de haleto tridimensionais estáveis são escolher composições com um fator de tolerância Goldschmidt entre 0,8 e 1,0 [43,44] ou um fator de tolerância Bartel abaixo de 4,18 [45], ambos essencialmente determinados pelos tamanhos dos íons envolvidos. Não surpreendentemente, é bem conhecido que os tamanhos dos cátions do sítio A afetam dramaticamente as propriedades das perovskitas de haleto [37,46-51]. Por exemplo, foi descoberto que aumentar o tamanho dos cátions substituintes em perovskitas baseadas em iodeto de chumbo e metilamônio ($MAPbI_3$) diminui a condutividade iônica e aumenta as energias de ativação [52]. O principal fator de diferenciação para cátions inorgânicos usados em perovskitas de haleto é frequentemente associado aos seus raios iônicos. No entanto, nos

cátions orgânicos usados em perovskitas de haleto, características diferentes do tamanho são de grande relevância, como composição química, existência de ligações duplas, possibilidade de formação de ligações de hidrogênio, forma e polaridade. Neste sentido, reconhece-se que aspectos além do tamanho dos cátions substituintes podem afetar as propriedades dos materiais resultantes [37,53-55].

Em trabalhos recentes, foi demonstrado que a formação de soluções sólidas monofásicas de perovskitas de haletos em composições do tipo $A_xMA_{1-x}PbI_3$ também é fortemente dependente do número de ligações N-H dos cátions substituintes A^+ [56]. Em termos de suas propriedades, foi sugerido que a formação de ligações de hidrogênio, também relacionadas às ligações N-H de cátions orgânicos, é um fator responsável pelo aumento da estabilidade no sistema $AC_xMA_{1-x}PbI_3$, dado que os cátions AC^+ possuem quatro ligações N-H, enquanto o MA^+ possui apenas três [57]. Entretanto, foi demonstrado que composições $GA_xMA_{1-x}PbI_3$ podem apresentar estabilidades térmicas menores que $MAPbI_3$ puro, embora o GA^+ possua seis ligações N-H [58,59]. Além disso, foi recentemente demonstrado que outras propriedades de interesse em perovskitas de haleto, por exemplo, permissividades e transições de fase em composições $GA_xFA_yMA_{1-x-y}PbI_3$, também não são dependentes somente do tamanho dos cátions substituintes [60]. Em relação às transições de fase, as expectativas baseadas no tamanho dos substituintes sugeririam que o sistema $GA_xMA_{1-x}PbI_3$ teria temperaturas de transição entre as fases tetragonal e cúbica menores do que aquelas do sistema $FA_xMA_{1-x}PbI_3$, dado que os cátions GA^+ são consideravelmente maiores do que do FA^+, mas é observado experimentalmente o oposto [58,60]. Em transições de baixa temperatura, foi demonstrado que a estabilização de outras simetrias pode depender de fatores como tamanho, número de ligações N-H e dipolos dos cátions substituintes [60]. A relevância da simetria estrutural, das transições de fase em perovskitas de haleto e efeitos na dinâmica de portadores de carga e aplicações optoeletrônicas foram recentemente assunto de artigos de revisão publicados em periódicos dentre os de maior fator de impacto na área de química, física e ciência dos materiais [61,62]. Por outro lado, os dipolos de cátions em perovskitas de haleto influenciam vários fatores, como separação de carga, estados de energia eletrônica, transições de fase, estrutura de defeito, transporte de íons e formação de domínios ferroicos [63-66]. Por exemplo, em comparação com $FA_xMA_{1-x}PbI_3$ e $GA_xMA_{1-x}PbI_3$ [60], as composições $EA_xMA_{1-x}PbI_3$ aparentemente têm permissividades dielétricas mais altas [67], o que pode ser devido ao fato de que a magnitude do dipolo intrínseco dos cátions EA^+ é maior do que a dos cátions FA^+ e GA^+. Ademais, mesmo características como as energias de *band*

gap [68], controladas majoritariamente pelas espécies nos sítios B e X, são indiretamente afetadas pelos cátions do sítio A [69]. Mostrou-se recentemente que os cátions FA^+ e GA^+ possuem efeitos inversos sobre a energia de *band gap* em perovskitas baseadas em $MAPbI_3$ [70] embora ambos os cátions levem a distorções estruturais similares, mas com diferentes intensidades. Por fim, em um trabalho recente, foi mostrado que os cátions EA^+, GA^+ e AC^+, que possuem raios muito similares, levam a resultados marcadamente distintos na estrutura, estabilidade e propriedades elétricas de perovskitas baseadas em $MAPbI_3$, o que foi relacionado aos seus números de ligações N-H e momentos de dipolo [71].

Considerando o exposto, resta claro que o tamanho, o momento dipolar e a capacidade de formar ligações de hidrogênio, que se relaciona com o número de ligações N-H, estão entre características importantes sob consideração ao lidar com cátions do sítio A em perovskitas de haleto. Ainda assim, uma sistematização clara de suas correlações e efeitos em propriedades específicas parece longe de ser alcançada. Desta forma, inspirando-se em uma perspectiva recente [51], se há uma relação entre as características dos cátions do sítio A e as propriedades das HPs, seria possível projetar uma combinação perfeita de cátions? Visando responder a essa questão, o presente projeto de pesquisa propôs estudar sistematicamente, com dados próprios e da literatura, quais, como e com que intensidade as características dos cátions do sítio A e suas misturas afetam as propriedades das perovskitas de haleto resultantes. Para tanto, obteve-se correlações diretas entre a composição química de cátions do sítio A e as propriedades dos materiais resultantes de forma a permitir um controle preditivo e, assim, a engenharia de cátions mistos do sítio A em vista de se atingir características finamente desenhadas para aplicações de interesse com perovskitas de haleto.

2. OBJETIVOS, HIPÓTESES E QUESTÕES A SEREM RESPONDIDAS

2.1 Objetivos

1) Sintetizar, por via mecanoquímica, pós de perovskitas de haleto contendo cátions mistos, e pastilhas policristalinas para estudar a estrutura, propriedades elétricas e optoeletrônicas intrínsecas e estabilidade dessas perovskitas;
2) Caracterizar a estrutura e composição de fases (simetria das estruturas das perovskitas formadas, parâmetros de rede, presença de fases secundárias e/ou de precursores não reagidos) usando difração de raios X, as transições de fase (temperaturas de transição) por calorimetria exploratória diferencial, a microestrutura (morfologia de grãos e porosidade) por microscopia eletrônica de varredura, as propriedades optoeletrônicas (energias de *band*

gap) por espectroscopia de refletância difusa de UV/Vis e as propriedades elétricas (condutividade elétrica, densidade de portadores, mobilidade de portadores e histerese corrente-tensão) por eletrometria I-V;

3) Estudar a cinética de degradação acelerada por luz acompanhando as mudanças estruturais e de composições de fases (via difração de raios X) e da microestrutura (via microscopia eletrônica de varredura) que ocorrem em função do tempo de exposição;
4) Coletar informações análogas na literatura, que permitam comparação direta com os dados coletados experimentalmente no presente trabalho, e montar um banco de dados amplo para servir de base a modelos estatísticos e computacionais capazes de correlacionar as características dos cátions e suas misturas no sítio A com as características das perovskitas resultantes.

2.2 Questão e hipóteses

- É possível obter propriedades finamente ajustáveis e prever as características de perovskitas de haleto de cátions mistos ainda não sintetizadas variando-se apenas os tipos e concentrações dos cátions do sítio A?

Para responder a essa questão, parte-se da hipótese de que dentro dos limites impostos pelos componentes presentes nos sítios B e X da perovskita, o domínio das propriedades das perovskitas de haleto se dá pelos cátions do sítio A. Conforme robustamente discutido na introdução, variações do raio iônico equivalente (r_A), número de ligações N-H (n_A) e momentos de dipolo (μ_A) dos cátions presentes afetam as propriedades das perovskitas. Nesse sentido, para conduzir o estudo, foram utilizadas composições químicas baseadas em $APbI_3$, em que A representa o conjunto de possíveis cátions do sítio A. Experimentalmente, foram usados os cátions MA^+, GA^+, FA^+, EA^+, AC^+ e TMS^+ (trimetilsulfônio), enquanto composições contendo os cátions referenciados na introdução e outros são encontradas na literatura. Por óbvio, para se estudar o efeito da composição do sítio A, os sítios B e X foram fixados e ocupados, respectivamente, pelo chumbo e iodo. Embora geralmente desfavorável por sua toxicidade, o uso do chumbo ainda é mais indicado para o sítio B em HPs aplicadas à energia [72], com a outra boa alternativa, o estanho, sendo muito propenso à oxidação [73]. Apesar dos potenciais problemas associados à toxicidade, estratégias são propostas para mitigar eventuais danos causados pelo chumbo [74]. Por sua vez, o uso do iodo se apoia no fato de que as energias de *band gap* das composições contendo outros haletos são maiores e potencialmente fora da faixa ótima para uso em células solares [75,76]. De qualquer forma, a

presente proposta trata de um estudo exploratório que servirá de base, com as devidas mudanças, para os casos em que outros íons estão presentes nos sítios B e X. Por fim, cita-se que as composições baseadas em $APbI_3$ são as mais presentes na literatura, o que permite a incorporação de maior quantidade de dados para serem usados nos tratamentos estatísticos e computacionais posteriores.

3. METODOLOGIA

A execução deste projeto de pesquisa será realizada em quatro etapas distintas, a saber, 1) síntese, 2) caracterização e extração das propriedades experimentais, 3) coleta bibliográfica e organização dos dados e 4) aplicação de modelos estatísticos e computacionais. A completa infraestrutura disponível para a produção e caracterização dos pós e das pastilhas policristalinas, além da experiência do grupo de pesquisa envolvido no tema proposto e dos colaboradores internos e externos à Instituição são aspectos favoráveis à execução da presente proposta. Os detalhes da metodologia e procedimentos são descritos a seguir.

3.1. Síntese das pastilhas policristalinas

O estudo experimental foi executado em 65 composições de cátions simples, duplos e triplos dentro do sistema $TMS_aAC_bEA_cGA_dFA_fMA_gPbI_3$. Inicialmente, os pós das composições foram preparados por rota mecanoquímica. Para cada composição, foram pesadas quantidades dos precursores iodeto de chumbo (PbI_2), de metilamônio (MAI), de formamidínio (FAI), de guanidínio (GAI), de acetamidínio (ACI), de etilamônio (EAI) e de trimetilsulfônio (TMSI) nas proporções desejadas. Os precursores foram misturados manualmente e moídos usando gral e pistilo de ágata natural por aproximadamente 50 min em condições ambientes. Os pós foram peneirados em peneira 250 μm (60 mesh) e depois aquecidos para homogeneização e remoção de umidade residual por ~ 3 horas a 60 °C. Com os pós de cada composição, foram feitas pastilhas pela adição de ~ 150 mg à cavidade circular de um molde metálico de 10 mm de diâmetro e posterior prensagem do pó com tensão de 250 MPa com auxílio de uma prensa hidráulica. Por fim, as pastilhas foram sinterizadas em forno elétrico a 120 °C por 3h.

3.2. Caracterização das amostras

As análises de difração de raios X (DRX) foram realizados em um difratômetro Rigaku Ultima IV com radiação Cu-Kα (λ = 1,5406 Å), sob 40 kV e 20 mA, em varredura

contínua (2° min^{-1}), com 2θ de 10 a 50°, à temperatura ambiente. As medidas de calorimetria exploratória diferencial (DSC) foram realizadas em um DSC 25 - TA Instruments de -90 a 100 °C, com taxa de aquecimento de 5 °C min^{-1}, sob atmosfera de nitrogênio. As imagens da microestrutura das amostras foram obtidas em microscópio eletrônico de varredura (MEV) Zeiss EVO LS15, com feixe de tungstênio, operando na faixa de tensão entre 10 e 20 kV, por análise de elétrons secundários. Os espectros de refletância difusa (UV/Vis) foram coletados em um espectrofotômetro Shimadzu UV-2600, na faixa de 200 a 1100 nm, no modo de baixa velocidade com intervalo de dados de 1 nm. Para as medidas elétricas, uma matriz de 3 x 3 eletrodos circulares de ouro de ~ 1,0 mm de diâmetro será depositada no topo das pastilhas usando uma máscara de cobre, enquanto um eletrodo de ouro inferior será depositado sobre toda a superfície da amostra. Dados de corrente em função do potencial (eletrometria I-V) foram coletados usando um eletrômetro Keithley 6517B. As medidas foram realizadas em condições ambiente, no escuro. Os dados foram coletados em ciclos de faixa de -5 a +5 V e vice-versa com uma taxa de varredura de 0,1 $V.s^{-1}$ e em potencial ascendente de 0 a +30 V com uma taxa de varredura de 0,5 $V.s^{-1}$. Ensaios de degradação acelerada por luz seriam conduzidos usando um sistema (caixa) fechado construído pelo próprio proponente e já empregado em trabalhos publicados [70,77]. As amostras seriam distribuídas uniformemente numa placa de Petri a ser posicionada no centro de uma plataforma giratória (6 Hz no sentido horário) e submetida à iluminação direta e indireta, pela reflexão por um espelho, da luz emitida por lâmpada LED. As especificações da lâmpada são potência de 6 W, fluxo luminoso de 560 lm e cor branco frio (6500 K). A caixa (30 x 20 x 15 cm) permanece fechada e selada com fita adesiva durante o teste. A umidade interna é mantida tão baixa quanto possível por béqueres preenchidos com sílica. A temperatura e a umidade relativa são monitoradas constantemente com um termo-higrômetro digital. A cada intervalo predefinido de 20 h até um total de 160 h, as amostras são removidas do sistema e dados de fases via difração de raios X seriam coletados no equipamento mencionado anteriormente, em varredura contínua com 2θ de 10 a 15°, à temperatura ambiente. Após cada medida, as amostras seriam novamente inseridas no sistema para continuidade do teste.

3.3 Coleta bibliográfica e organização dos dados

Os dados experimentais coletados no presente trabalho foram divididos em duas categorias. Na primeira categoria, chamada de PHASE.DATA, foram agrupadas informações sobre 1) as composições formarem ou não perovskitas no sentido *stricto sensu*, 2) se são

sistemas monofásicos ou multifásicos (isto é, se em sistemas mistos há formação de fases secundárias decorrentes de se estar acima dos limites de solubilidade), 3) qual a simetria (tetragonal, cúbica, ortorrômbica etc.), na temperatura ambiente, dos sistemas monofásicos e 4) as temperaturas de transição de fases. Nessa categoria, foram incluídos também os dados da literatura nos diferentes sistemas simples e mistos. O proponente revisou cerca de 200 artigos que continham potenciais informações de fases dentro dos critérios 1 a 4. Na soma entre os resultados experimentais do presente trabalho acrescidos dos dados da literatura, foram agrupadas informação de fases de mais de 200 diferentes composições de cátions mistos do sítio A. Na segunda categoria, chamada de PROP.DATA, os dados experimentais de energia de *band gap* e propriedades de transporte elétrico coletados no presente trabalho foram organizados apenas para as composições monofásicas obtidas anteriormente.

3.4 Aplicação de modelos estatísticos e computacionais

A generalização do método, de forma a permitir a correlação entre composição de cátions mistos e as características das perovskitas resultantes e, consequentemente, a possibilidade preditiva de ajuste racional de componentes do sítio A, exige que as composições dos diferentes cátions sejam convertidas em descritores. Conforme mencionado, os descritores mais simples são o raio iônico equivalente (r_A), número de ligações N-H (n_A) e momentos de dipolo (μ_A). Para cada cátion do sítio A, esses descritores podem ser estimados, respectivamente, pelos valores tabelados e/ou métodos descritos em [44,78,79], pelas estruturas moleculares e por cálculos simplificados feitos via *MolCalc* [80]. Para uma composição de cátions do sítio A, supondo um sistema arbitrário $A_\alpha PbI_3$, tal que A_α representa uma mistura $TMS_aAC_bEA_cGA_dFA_fMA_gK_hRb_iCs_jIM_kHy_lMHy_mAZi_nAZe_oDMA_p$... ($\sum_{\alpha=a}^{\infty}\alpha = 1$), os descritores efetivos para o raio iônico equivalente, número de ligações N-H e momentos de dipolo foram dados, respectivamente, por $r_{eff} = \sum_{\alpha=a}^{\infty}\alpha r_A$, $n_{eff} = \sum_{\alpha=a}^{\infty}\alpha n_A$ e $\mu_{eff} = \sum_{\alpha=a}^{\infty}\alpha\mu_A$. Por fim, para descrever conjuntos de cátions com inevitavelmente diferentes graus de complexidade, será usado um quarto descritor, a saber, a entropia configuracional de mistura (ΔS_c), que, para mudanças apenas no sítio A em uma rede do tipo perovskita de composição genérica ABX_3, é dada por $\Delta S_c = -\frac{R}{5}\sum_{\alpha=a}^{\infty}\alpha(\ln\alpha)$. Tomado esse cenário, os dados das categorias PHASE.DATA e PROP.DATA foram todos convertidos nesses quatro descritores. Nessa forma, os dados foram usados em conjunto com técnicas de aprendizado de máquina (*machine learning*, ML) para identificação de padrões nas relações composição-propriedades e possibilitar a previsão de propriedades de

composições que não tenham sido sintetizadas. Técnicas de ML são amplamente empregadas em química aplicada e ciência dos materiais [82-85] e, especificamente, também em perovskitas de haleto [86-88]. Em específico, para a categoria PHASE.DATA, espera-se utilizar a abordagem SISSO (*Sure Independence Screening and Sparsifying Operator*) [88,89] que combina regressão simbólica e detecção comprimida para identificar modelos e descritores em expressões matemáticas explícitas para ciência de materiais e que já foi usada com sucesso na previsão de formação de perovskitas em compostos ternários (ou seja, não incluindo soluções sólidas) [45]. Essa abordagem deve fornecer estimativas quantitativas sobre os critérios 1 a 4 descritos na subseção anterior. Por sua vez, para os dados da categoria PROP.DATA, visa-se sobretudo encontrar correlações de quais descritores são importantes, como eles afetam e quão intensos são seus efeitos sobre as propriedades estudadas. Embora, *a priori*, o número de dados seja relativamente limitado, espera-se utilizar métodos, no contexto do aprendizado de máquina, mais especificamente, as redes neurais da família ART (*adaptive resonance theory*) [90], *viz.*, a arquitetura ARTMAP-*Fuzzy*. A grande vantagem do uso desta família de redes neurais artificiais (ANN) é observada a partir das suas características intrínsecas que são: estabilidade, plasticidade, simplicidade e rapidez de treinamento. A estabilidade representa a garantia plena de convergência na fase de treinamento da ANN. A plasticidade, uma qualidade promissora, pois permite realizar o treinamento de forma contínua (incremental), ou seja, trata-se de uma ANN enquanto realiza o diagnóstico, permite incluir novos conhecimentos sem a necessidade de reinicializar todo o processo de treinamento, esquivando-se da conhecida condição cognitiva idealizada e caracterizada como vazio completo (“tábula rasa”, segundo Aristóteles). Guardada as justas particularidades, neste caso, tem-se uma execução similar à ação humana, a qual é aprender continuamente. Portanto, trata-se da idealização do aprendizado contínuo [91] aplicado conforme novos dados experimentais forem coletados pelo grupo, nas mesmas condições descritas na subseção 3.2, e inseridos no banco de dados dessa categoria. A depender dos resultados, também poderão ser usadas outras ANN a serem discutidas posteriormente com os colaboradores do grupo.

4. CRONOGRAMA

O presente projeto de pesquisa teve a previsão de duração de 24 meses (4 semestres), com início em março de 2025. O cronograma das atividades é resumido na tabela a seguir.

ATIVIDADE	SEMESTRE			
	1	2	3	4
1. Síntese das pastilhas policristalinas	X	X		
2. Caracterização das amostras		X	X	
3. Coleta bibliográfica e organização dos dados		X	X	
4. Aplicação de modelos estatísticos e computacionais			X	X
5. Redação de artigos e do relatório final			X	X

Oportunamente, o proponente do presente projeto de pesquisa participaria de congressos e apresentaria os resultados em eventos científicos na área de estudo. Estava prevista colaboração com demais projetos desenvolvidos no grupo, coorientação de alunos em nível de iniciação científica, mestrado e doutorado, e ministração de aulas de disciplinas do departamento aos cursos de graduação e pós-graduação do campus.

5. RESULTADOS E DISCUSSÃO

A análise qualitativa entre composição-estrutura-propriedades envolvendo os dados experimentais coletados para este trabalho foram organizados em um artigo atualmente em revisão na revista Physica Scripta. A versão submetida do trabalho, intitulado "*A-site composition-property relations in halide perovskites: which descriptors do what?*" encontra-se em anexo.

Por sua vez, o estudo quantitativo visando a previsão das temperaturas de transição de fases em composições quaisquer de perovskitas do tipo $APbI_3$, está organizado no trabalho intitulado "*Empirical equations to estimate tetragonal to cubic phase transitions in lead iodide perovskites*", também em anexo, e será imediatamente submetido a uma revista de alto impacto tão logo o anterior for aceito.

Finalmente, um outro estudo quantitativo para previsão de energias de *band gap* em perovskitas de haleto de composição qualquer está atualmente concluído. Esse trabalho, provisoriamente intitulado "*Empirical equations to estimate band gap energies in halide perovskites*" será escrito em breve e submetido a uma revista de alto impacto tão logo o anterior for aceito. Informações prévias também estão disponíveis em anexo.

6. PUBLICAÇÕES DERIVADAS DO PROJETO

Durante a vigência deste projeto, os seguintes trabalhos envolvendo o pesquisador foram publicados:

- R.K. Mishra, F.B. Minussi, P. Kumari, R.R. Shahi, R.P. Yadav, F. Emmerling, and E.B. Araújo, "Effect of sintering temperature on structural, microstructural, and dielectric properties of (Co0.2Fe0.2Ni0.2Al0.2Ti0.2)3O4 multi-cations high entropy oxides," Ceramics International **51**(27, Part C), 55341–55354 (2025).
- F.B. Minussi, R.C.D.O. Santos, M.A.D.M. Teixeira, R.M. Silva, and E.B. Araújo, "An Exploration of the Role of Acetamidinium Substitution in Methylammonium Lead Iodide Perovskites," ChemPhysChem **27**(1), e202500259 (2026).
- M.A.M. Teixeira, K.N. Romanyuk, F.B. Minussi, V.V. Shvartsman, A.L. Kholkin, D.C. Lupascu, S. Salamon, H. Wende, J.F. Carvalho, J.A. Eiras, and E.B. Araújo, "Pathways to obtain high-entropy BiFeO3-based thin films modified by rare earth cations," Ceramics International, (2026).

7. OUTRAS PUBLICAÇÕES

Durante o período do presente projeto, o pesquisador atuou como coorientador em um projetos de iniciação científica, e um projeto de mestrado, listados abaixo:

- Rogério Marcos da Silva Jr. *Triagem da formação e estabilidade de fases de perovskitas de haleto contendo cátions orgânicos mistos* (iniciação científica), início em 08/2024 e finalizado em 07/2025.
- Rafaela Coutinho de Oliveira Santos. *Descrição físico-química de perovskitas do sistema ACxMA1−xPbI3* (mestrado em Ciência dos Materiais), início em 03/2024 e finalizado em 12/2025.

8. DISCIPLINAS MINISTRADAS

O pesquisador ministrou a disciplina "Tópicos especiais: Introdução aos materiais cerâmicos" no Programa de Pós-Graduação em Ciência dos Materiais durante o primeiro semestre do ano letivo de 2025.

9. COMENTÁRIOS DO SUPERVISOR

O projeto de pós-doutorado do Dr. Fernando Brondani Minussi, com início em 01/03/2025 e término previsto para 28/02/2027, conforme estabelecido no cronograma do plano de trabalho inicial, foi conduzido e executado com sucesso ao longo do ano de 2025. Entretanto, como resultado de aprovação em concurso público, no dia 10/03/2026 o Dr. Fernando Brondani Minussi tomou posse como professor adjunto da Universidade Estadual de Santa Catarina (UDESC), em Joinville, motivo pelo qual solicita-se a finalização antecipada do projeto em 09/03/2026. O relatório apresentado foi muito bem redigido e descreveu adequadamente os principais resultados obtidos no período (01/03/2025 à 09/03/2025), demonstrando que, apesar do projeto ter sido executado na metade do tempo estabelecido, a grande maioria dos objetivos foram cumpridos satisfatoriamente e com desenvoltura, com foco na exploração de perovskitas de haleto contendo cátions mistos e na cinética de degradação induzida por iluminação. Considerando o dinamismo da execução do projeto, traduzido em resultados publicados em artigos científicos em periódicos internacionais de bom fator de impacto, não há dúvidas de que os estudos promoveram avanços relevantes na área do conhecimento e contribuíram para a consolidação da linha de pesquisa do Prof. Fernando. Considero aprovado o relatório científico apresentado, parabenizo o Prof. Fernando Minussi pelo excelente trabalho realizado e desejo sucesso na carreira que se inicia.

10. REFERÊNCIAS

[1] Q. A. Akkerman and L. Manna, *ACS Energy Lett.*, 2020, **5**, 604–610.

[2] J. Yan, H. Li, M. H. Aldamasy, C. Frasca, A. Abate, K. Zhao and Y. Hu, *Chem. Mater.*, 2023, **35**, 2683–2712.

[3] Y. Cho, H. R. Jung and W. Jo, *Nanoscale*, 2022, **14**, 9248–9277.

[4] L. Chouhan, S. Ghimire, C. Subrahmanyam, T. Miyasaka and V. Biju, *Chem. Soc. Rev.*, 2020, **49**, 2869–2885.

[5] A. Dey, J. Ye, A. De, E. Debroye, S. K. Ha, E. Bladt, A. S. Kshirsagar, Z. Wang, J. Yin, Y. Wang, L. N. Quan, F. Yan, M. Gao, X. Li, J. Shamsi, T. Debnath, M. Cao, M. A. Scheel, S. Kumar, J. A. Steele, M. Gerhard, L. Chouhan, K. Xu, X. Wu, Y. Li, Y. Zhang, A. Dutta, C. Han, I. Vincon, A. L. Rogach, A. Nag, A. Samanta, B. A. Korgel, C.-J. Shih, D. R. Gamelin, D. H. Son, H. Zeng, H. Zhong, H. Sun, H. V. Demir, I. G. Scheblykin, I. Mora-Seró, J. K. Stolarczyk, J. Z. Zhang, J. Feldmann, J. Hofkens, J. M. Luther, J. Pérez-Prieto, L. Li, L. Manna, M. I. Bodnarchuk, M. V. Kovalenko, M. B. J. Roeffaers,

N. Pradhan, O. F. Mohammed, O. M. Bakr, P. Yang, P. Müller-Buschbaum, P. V. Kamat, Q. Bao, Q. Zhang, R. Krahne, R. E. Galian, S. D. Stranks, S. Bals, V. Biju, W. A. Tisdale, Y. Yan, R. L. Z. Hoye and L. Polavarapu, *ACS Nano*, 2021, **15**, 10775–10981.

[6] A. Gaurav, A. Das, A. Paul, A. Jain, B. D. Boruah and M. Abdi-Jalebi, *Journal of Energy Storage*, 2024, **88**, 111468.

[7] W. A. Dunlap-Shohl, Y. Zhou, N. P. Padture and D. B. Mitzi, *Chem. Rev.*, 2019, **119**, 3193–3295.

[8] L. Zhang, L. Mei, K. Wang, Y. Lv, S. Zhang, Y. Lian, X. Liu, Z. Ma, G. Xiao, Q. Liu, S. Zhai, S. Zhang, G. Liu, L. Yuan, B. Guo, Z. Chen, K. Wei, A. Liu, S. Yue, G. Niu, X. Pan, J. Sun, Y. Hua, W.-Q. Wu, D. Di, B. Zhao, J. Tian, Z. Wang, Y. Yang, L. Chu, M. Yuan, H. Zeng, H.-L. Yip, K. Yan, W. Xu, L. Zhu, W. Zhang, G. Xing, F. Gao and L. Ding, *Nano-Micro Lett.*, 2023, **15**, 177.

[9] A. K. Jena, A. Kulkarni and T. Miyasaka, *Chem. Rev.*, 2019, **119**, 3036–3103.

[10] M. Cheng, J. Jiang, C. Yan, Y. Lin, M. Mortazavi, A. B. Kaul and Q. Jiang, *Nanomaterials*, 2024, **14**, 391.

[11] F. Temerov, Y. Baghdadi, E. Rattner and S. Eslava, *ACS Appl. Energy Mater.*, 2022, **5**, 14605–14637.

[12] H. Guthrey and J. Moseley, *Advanced Energy Materials*, 2020, **10**, 1903840.

[13] S. Narayanan, N. Parikh, M. M. Tavakoli, M. Pandey, M. Kumar, A. Kalam, S. Trivedi, D. Prochowicz and P. Yadav, *Eur J Inorg Chem*, 2021, **2021**, 1201–1212.

[14] K. Lê, N. Heshmati and S. Mathur, *Nano Convergence*, 2023, **10**, 47.

[15] Y. Wu, J. Feng, Z. Yang, Y. Liu and S. (Frank) Liu, *Advanced Science*, 2023, **10**, 2205536.

[16] L. Romani and L. Malavasi, *ACS Omega*, 2020, **5**, 25511–25519.

[17] P. V. Shinde, A. Patra and C. S. Rout, *J. Mater. Chem. C*, 2022, **10**, 10196–10223.

[18] H. Kim, J. S. Han, J. Choi, S. Y. Kim and H. W. Jang, *Small Methods*, 2018, **2**, 1700310.

[19] Q. Liu, Y. Yuan, J. Liu, W. Wang, J. Chen and W. Xu, *Materials Today Electronics*, 2024, **8**, 100099.

[20] Y. Zhou, J. Wang, D. Luo, D. Hu, Y. Min and Q. Xue, *Nano Energy*, 2022, **94**, 106949.

[21] Best Research-Cell Efficiency Chart, https://www.nrel.gov/pv/cell-efficiency.html (último acesso: 23 dez. 2024).

[22] P. Zhu, C. Chen, J. Dai, Y. Zhang, R. Mao, S. Chen, J. Huang and J. Zhu, *Advanced Materials*, 2024, **36**, 2307357.

[23] Q. Wei, D. Zheng, L. Liu, J. Liu, M. Du, L. Peng, K. Wang and S. Liu, *Advanced Materials*, 2024, 2406295.

[24] B. Li, S. Li, J. Gong, X. Wu, Z. Li, D. Gao, D. Zhao, C. Zhang, Y. Wang and Z. Zhu, *Chem*, 2024, **10**, 35–47.

[25] J. Zhuang, J. Wang and F. Yan, *Nano-Micro Lett.*, 2023, **15**, 84.

[26] H. Zhu, S. Teale, M. N. Lintangpradipto, S. Mahesh, B. Chen, M. D. McGehee, E. H. Sargent and O. M. Bakr, *Nat Rev Mater*, 2023, **8**, 569–586.

[27] N. Ahn and M. Choi, *Advanced Science*, 2024, **11**, 2306110.

[28] K. Liu, S. Rafique, S. F. Musolino, Z. Cai, F. Liu, X. Li, Y. Yuan, Q. Bao, Y. Yang, J. Chu, X. Peng, C. Nie, W. Yuan, S. Zhang, J. Wang, Y. Pan, H. Zhang, X. Cai, Z. Shi, C. Li, H. Wang, L. Deng, T. Hu, Y. Wang, Y. Wang, S. Chen, L. Shi, P. Ayala, J. E. Wulff, A. Yu and Y. Zhan, *Joule*, 2023, **7**, 1033–1050.

[29] Y. Wang, Y. Cheng, C. Yin, J. Zhang, J. You, J. Wang, J. Wang and J. Zhang, *Nano-Micro Lett.*, 2024, **16**, 183.

[30] H. Chen, L. Chu and W. Yan, *Adv Funct Materials*, 2024, 2412389.

[31] A. N. Iqbal, K. W. P. Orr, S. Nagane, J. Ferrer Orri, T. A. S. Doherty, Y. Jung, Y. Chiang, T. A. Selby, Y. Lu, A. J. Mirabelli, A. Baldwin, Z. Y. Ooi, Q. Gu, M. Anaya and S. D. Stranks, *Advanced Materials*, 2024, **36**, 2307508.

[32] M. Byranvand, C. Otero-Martínez, J. Ye, W. Zuo, L. Manna, M. Saliba, R. L. Z. Hoye and L. Polavarapu, *Advanced Optical Materials*, 2022, **10**, 2200423.

[33] D. Ghosh, A. R. Smith, A. B. Walker and M. S. Islam, *Chem. Mater.*, 2018, **30**, 5194–5204.

[34] J. Hieulle, D. Son, A. Jamshaid, X. Meng, C. Stecker, R. Ohmann, Z. Liu, L. K. Ono and Y. Qi, *Adv Funct Materials*, 2023, **33**, 2211097.

[35] Y. Li, Y. Wang, Z. Xu, B. Peng and X. Li, *ACS Nano*, 2024, **18**, 10688–10725.

[36] M. M. Byranvand, C. Otero-Martínez, J. Ye, W. Zuo, L. Manna, M. Saliba, R. L. Z. Hoye and L. Polavarapu, *Advanced Optical Materials*, 2022, **10**, 2200423.

[37] J.-W. Lee, S. Tan, S. I. Seok, Y. Yang and N.-G. Park, *Science*, 2022, **375**, eabj1186.

[38] J. Tian, D. B. Cordes, A. M. Z. Slawin, E. Zysman-Colman and F. D. Morrison, *Inorg. Chem.*, 2021, **60**, 12247–12254.

[39] M. Mączka, M. Ptak, K. Fedoruk, D. Stefańska, A. Gągor, J. K. Zaręba and A. Sieradzki, *J. Mater. Chem. C*, 2024, **12**, 1396–1405.

[40] M. Mączka, M. Ptak, A. Gągor, J. K. Zaręba, X. Liang, S. Balčiūnas, O. A. Semenikhin, O. I. Kucheriv, I. A. Gural'skiy, S. Shova, A. Walsh, J. Banys and M. Šimėnas, *Chem. Mater.*, 2023, **35**, 9725–9738.

[41] Q. Wang, F. Lin, C.-C. Chueh, T. Zhao, M. Eslamian and A. K.-Y. Jen, *Materials Today Energy*, 2018, **7**, 161–168.

[42] J. Gebhardt and A. M. Rappe, *Advanced Materials*, 2019, **31**, 1802697.

[43] G. Kieslich, S. Sun and A. K. Cheetham, *Chem. Sci.*, 2014, **5**, 4712–4715.

[44] G. Kieslich, S. Sun and A. K. Cheetham, *Chem. Sci.*, 2015, **6**, 3430–3433.

[45] C. J. Bartel, C. Sutton, B. R. Goldsmith, R. Ouyang, C. B. Musgrave, L. M. Ghiringhelli and M. Scheffler, *Sci. Adv.*, 2019, **5**, eaav0693.

[46] H. B. Lee, N. Kumar, B. Tyagi, S. He, R. Sahani and J.-W. Kang, *Materials Today Energy*, 2021, **21**, 100759.

[47] H. Park, S. Kumar, S. Chawla and F. El-Mellouhi, *Molecules*, 2021, **26**, 6184.

[48] H. Park, A. Ali, R. Mall, H. Bensmail, S. Sanvito and F. El-Mellouhi, *Mach. Learn.: Sci. Technol.*, 2021, **2**, 025030.

[49] F. Ünlü, E. Jung, J. Haddad, A. Kulkarni, S. Öz, H. Choi, T. Fischer, S. Chakraborty, T. Kirchartz and S. Mathur, *APL Materials*, 2020, **8**, 070901.

[50] C. Wu, D. Guo, P. Li, S. Wang, A. Liu and F. Wu, *Phys. Chem. Chem. Phys.*, 2020, **22**, 3105–3111.

[51] S. Jin, *ACS Energy Lett.*, 2021, **6**, 3386–3389.

[52] D. W. Ferdani, S. R. Pering, D. Ghosh, P. Kubiak, A. B. Walker, S. E. Lewis, A. L. Johnson, P. J. Baker, M. S. Islam and P. J. Cameron, *Energy Environ. Sci.*, 2019, **12**, 2264–2272.

[53] Z. Zhang, M. Li, R. Li, X. Zhuang, C. Wang, X. Shang, D. He, J. Chen and C. Chen, *Advanced Materials*, 2024, **36**, 2313860.

[54] M. Saski, S. Sobczak, P. Ratajczyk, M. Terlecki, W. Marynowski, A. Borkenhagen, I. Justyniak, A. Katrusiak and J. Lewiński, *Small*, 2024, 2403685.

[55] P. Singh, D. R. Ceratti, Y. Soffer, S. Bera, Y. Feldman, M. Elbaum, D. Oron, D. Cahen and G. Hodes, *J. Phys. Chem. C*, 2024, **128**, 19999–20008.

[56] F. B. Minussi, R. M. da Silva and E. B. Araújo, *Inorg. Chem. Front.*, 2023, **10**, 6129–6133.

[57] P. Singh, R. Mukherjee and S. Avasthi, *ACS Appl. Mater. Interfaces*, 2020, **12**, 13982–13987.

[58] F. B. Minussi, R. M. Da Silva and E. B. Araújo, *J. Phys. Chem. C*, 2023, **127**, 8814–8824.

[59] F. B. Minussi, R. M. Silva, J. F. Carvalho and E. B. Araújo, *J. Mater. Chem. C*, 2024, **12**, 5138–5149.

[60] F. B. Minussi, R. M. Silva and E. B. Araújo, *Small*, 2024, **20**, 2305054.

[61] M. Simenas, A. Gagor, J. Banys and M. Maczka, *Chem. Rev.*, 2024, **124**, 2281–2326.

[62] H. Guo, S. Yi, S. Yang, Y. Jiao, S. Qin, R. Li, Y. Wang, H. Li, Y. Xia, Y. Zhang, Z. Liang, R. Yan, H. Liu and J. Zhao, *Adv Funct Materials*, 2023, **33**, 2214180.

[63] W. Geng, Q. Hu, C. Tong, Z. Tang and L. Liu, *Advcd Theory and Sims*, 2019, **2**, 1900041.

[64] S. Maheshwari, S. Patwardhan, G. C. Schatz, N. Renaud and F. C. Grozema, *Phys. Chem. Chem. Phys.*, 2019, **21**, 16564–16572.

[65] S. Maheshwari, M. B. Fridriksson, S. Seal, J. Meyer and F. C. Grozema, *J. Phys. Chem. C*, 2019, **123**, 14652–14661.

[66] M. Pazoki, M. J. Wolf, T. Edvinsson and J. Kullgren, *Nano Energy*, 2017, **38**, 537–543.

[67] M. Mączka, A. Gagor, J. K. Zaręba, D. Stefanska, M. Drozd, S. Balciunas, M. Šimėnas, J. Banys and A. Sieradzki, *Chem. Mater.*, 2020, **32**, 4072–4082.

[68] A. Walsh, *J. Phys. Chem. C*, 2015, **119**, 5755–5760.

[69] M. P. Hautzinger, W. Mihalyi-Koch and S. Jin, *Chem. Mater.*, 2024, **36**, 10408–10420.

[70] F. B. Minussi, R. M. Silva and E. B. Araújo, *Applied Physics Letters*, 2024, **124**, 153901.

[71] F. B. Minussi, R. M. Silva, J. C. S. Moraes and E. B. Araújo, *Phys. Chem. Chem. Phys.*, 2024, **26**, 20770–20784.

[72] T. A. Berhe, W.-N. Su and B. J. Hwang, *Inorganics*, 2024, **12**, 182.

[73] J. Pascual, G. Nasti, M. H. Aldamasy, J. A. Smith, M. Flatken, N. Phung, D. D. Girolamo, S.-H. Turren-Cruz, M. Li, A. Dallmann, R. Avolio and A. Abate, *Mater. Adv.*, 2020, **1**, 1066–1070.

[74] X. Wang, B. Dong, M. Feng, D.-J. Xue and S.-M. Wang, *J. Mater. Chem. A*, 2022, **10**, 15861–15864.

[75] M. H. Miah, M. U. Khandaker, M. B. Rahman, M. Nur-E-Alam and M. A. Islam, *RSC Adv.*, 2024, **14**, 15876–15906.

[76] C. P. Muzzillo, C. V. Ciobanu and D. T. Moore, *Mater. Horiz.*, 2024, **11**, 3662–3694

[77] F. B. Minussi and E. B. Araújo, *Phys. Chem. Chem. Phys.*, 2024, **26**, 26479–26488.

[78] R. D. Shannon, *Acta Cryst A*, 1976, **32**, 751–767.

[79] W. Travis, E. N. K. Glover, H. Bronstein, D. O. Scanlon and R. G. Palgrave, *Chem. Sci.*, 2016, **7**, 4548–4556.

[80] J. H. Jensen and J. C. Kromann, *J. Chem. Educ.*, 2013, **90**, 1093–1095.

[81] F. Oviedo, J. L. Ferres, T. Buonassisi and K. T. Butler, *Acc. Mater. Res.*, 2022, **3**, 597–607.

[82] A. Y.-T. Wang, R. J. Murdock, S. K. Kauwe, A. O. Oliynyk, A. Gurlo, J. Brgoch, K. A. Persson and T. D. Sparks, *Chem. Mater.*, 2020, **32**, 4954–4965.

[83] J. Wei, X. Chu, X. Sun, K. Xu, H. Deng, J. Chen, Z. Wei and M. Lei, *InfoMat*, 2019, **1**, 338–358.

[84] X. Zhong, B. Gallagher, S. Liu, B. Kailkhura, A. Hiszpanski and T. Y.-J. Han, *npj Comput Mater*, 2022, **8**, 204.

[85] A. R. Hering, M. Dubey and M. S. Leite, *APL Energy*, 2023, **1**, 020901.

[86] J. Liang, T. Wu2, Z. Wang, Y. Yu, L. Hu, H. Li, X. Zhang, X. Zhu and Y. Zhao, *Energy Mater*, 2022, **2**, 200016.

[87] N. Parikh, M. Karamta, N. Yadav, M. Mahdi Tavakoli, D. Prochowicz, S. Akin, A. Kalam, S. Satapathi and P. Yadav, *Journal of Energy Chemistry*, 2022, **66**, 74–90.

[88] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler and L. M. Ghiringhelli, *Phys. Rev. Materials*, 2018, **2**, 083802.

[89] T. A. R. Purcell, M. Scheffler and L. M. Ghiringhelli, *The Journal of Chemical Physics*, 2023, **159**, 114110.

[90] S. Grossberg, *Neural Networks*, 2013, **37**, 1–47.

[91] M. A. Da Silva, T. Abreu, C. R. Santos-Júnior and C. R. Minussi, *Electric Power Systems Research*, 2021, **201**, 107545.

A-site composition-property relations in halide perovskites: which descriptors do what?

Fernando B. Minussi[1,2*#], Rogerio M. Silva Jr.[3#], Marco Antonio M. Teixeira[1], Ricardo P. Bonini[2], Thays A. Abreu[4], Jose A. Eiras[2], Eudes B. Araujo[1]

[1] Department of Physics and Chemistry, São Paulo State University, Ilha Solteira, SP, Brazil

[2] Department of Physics, Federal University of São Carlos, São Carlos, SP, Brazil

[3] Department of Electrical Engineering, São Paulo State University, Ilha Solteira, SP, Brazil

[4] Department of Mathematics, São Paulo State University, Ilha Solteira, SP, Brazil

Abstract

The vastness of halide perovskite (HP) properties and applications is greatly a consequence of their compositional versatility. A large part of this versatility stems from A-site compositional engineering in the ABX_3 stoichiometry, enabling fine-tuning of the desired properties. The current literature provides in-depth knowledge of how A-site cations affect these properties. However, there is a certain sense of mystery surrounding understanding exactly why the different cations cause different effects. To provide more insight into finding answers to that, we here propose describing diverse A-site HP compositions using a combination of their effective descriptors, namely the radii, dipole moments, number of N-H bonds, and entropy of mixing. For this purpose, a total of 65 three-dimensional ABX_3 compositions were synthesized and their structures, optoelectronic, and electrical behavior were characterized. The extracted properties of the obtained single-phase solid solutions, namely, tetragonal-to-cubic transition temperature, band gap, Urbach energies, electrical conductivity, and I-V hysteresis, were correlated with the four descriptors using statistical tools. Examples of relevant effects were found for the entropy in reducing and number of N-H bonds in increasing the phase transition temperature, the dipole moment in increasing the band gap, the radius in reducing the Urbach energy and the dipole moment and N-H bonds in mitigating the charge transport characteristics. The results provide suitable rules of thumb for qualitative prediction of property changes upon varying the A-site in 3D ABX_3 halide perovskites.

Keywords: Halide perovskites, A-site cations, structural descriptors, composition-property relations, phase transitions, optoelectronic properties, electrical properties.

* fbminussi@gmail.com / f.minussi@unesp.br

# Authors contributed equally

1. INTRODUCTION

Halide perovskites (HPs) are a relatively recent but well-established, consolidated class of materials that are now finding applications ranging from solar cells [1], lasers [2], radiation detectors [3], and memristors and neuromorphic devices [4] to batteries [5-7], gas sensing of volatile organic compounds [8,9], and photocatalysis for H_2 evolution [10] and CO_2 reduction [11]. This versatility largely stems from the immense range of compositions, structures, and synthetic approaches of these materials, which enable a wide range of chemical and physical properties.

Adhering to a strict definition of a perovskite [12,13], HPs have a stoichiometric formula based on ABX_3, where A is commonly a monovalent cation, such as Rb^+, Cs^+, MA^+ (methylammonium), FA^+ (formamidinium), AZR^+ (aziridinium), and MHy^+ (methylhydrazinium), B a divalent cation, mainly Pb^{2+} and Sn^{2+}, and X a halide anion, commonly Cl^-, Br^-, and I^- [14,15], constructed by the three-dimensional (3D) corner-sharing connection of BX_6 octahedra. Despite this, perovskite is also used to describe other structures with diverse arrangements of BX_6 octahedra. Some examples of these low-dimensional (LD) structures are the Dion-Jacobsson (DJ), and Ruddlesden-Popper (RP) structures [16-19].

Changing any of the sites in the perovskite structure is possible and leads to property customization of the resulting materials. As such, altering the composition of A-site cations is a common strategy for tailoring the properties of perovskite materials and simultaneously increasing their stability [20]. Within the diverse compositions of LD structures, a plethora of A-site (a.k.a. the spacer cation) species can be used, which cause extensive changes in the structure and properties of the compounds [21,22]. In turn, in actual HPs, the constraints on the formation of three-dimensional connectivity greatly limit which cations can be used in single-A-site compounds. However, sensible changes can be made by partially incorporating cations in solid solution mixtures that alone do not normally form 3D perovskites, such as GA^+ (guanidinium), DMA^+ (dimethylammonium), AC^+ (acetamidinium), IM^+ (imidazolium), and EA^+ (ethylammonium) [15,23].

The establishment of which and how the characteristics of A-site cations relate to the properties of HPs has been actively discussed [15,23-28]. Although this direct determination would allow a rational compositional design of HPs and their devices, it has arguably not yet been achieved. One representative example is the band gap energy of these materials. A couple of years ago, it was found that the band gap increases as the ionic radius of the A-site decreases, leading to modulation of the Pb–I bond from size effects in the A-site cations [29,30]. On the

other hand, a bowing trend was observed in a recent review regarding the A-site cation chemistry in halide perovskites [27]. However, in mixed cation systems, the behavior can be even more intricate, as seen in the $GA_xFA_yMA_{1-x-y}$ system, in which FA^+ reduces the band gap, while GA^+ increases it, even though both are larger than MA^+ [31]. A similar behavior is observed in the $Cs_xFA_yMA_{1-x-y}$ system, where Cs^+ also increases the band gap, but, in contrast, it is smaller than MA^+ [32]. The transition temperatures between the tetragonal and cubic phases are another example. While it can be argued that using larger cations in $MAPbI_3$-based perovskites brings them closer to the cubic phase, the trends cannot be explained solely by cation radius [33,34]. Recognizing these experimental facts, we recently studied how cations of nearly equal size (namely EA^+, GA^+, and AC^+) produce markedly different effects when used as substituents in $MAPbI_3$-based perovskites [35]. Essentially, the current literature provides in-depth information on how cations affect perovskite properties, but it appears we do not yet know exactly why the effects differ from one cation to another.

In this scenario, heavily motivated by a prominent perspective [36], we aim to add a bit more seasoning to the question: can we find the perfect A-cations for halide perovskites? To discuss this, we ask ourselves which cation(s) would be ideal for a given intention or application. Maybe it would be important to have the largest electrical conductivity or the lowest band gap. So, how can desired properties be achieved? Is it possible to obtain them through A-site engineering? Which aspects of the A-site cation or mixture of cations are important to control these properties?

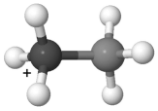
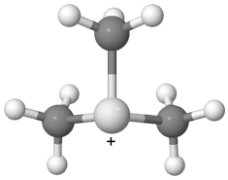
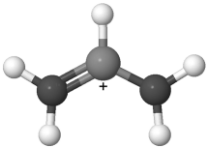
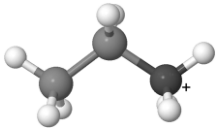
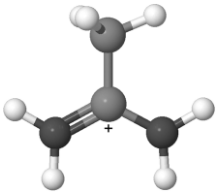
Here, we begin with the hypothesis that it is possible to describe the properties of halide perovskites by the characteristics of the organic cations used in the A-site. To test this hypothesis, we fixed B = Pb^{2+} and X = I^-, and changed only the A-site. To correlate which and how the characteristics of the A-site cations affect the properties of the perovskites formed, the A-site compositions are transformed into four descriptors. The first descriptor used is obviously the cation size, which is the most relevant characteristic considered to change and explain the properties of halide perovskites. The second descriptor is the number of N-H bonds, taken as a measure of the capability of forming hydrogen bonds between the cations and the iodide ions of the inorganic framework. Despite some statements to the contrary [37], hydrogen bonds are generally well-recognized to play a major role in halide perovskites [38]. Third, we used the intrinsic dipole moment of the cations, known to be relevant on both LD and 3D perovskites [15,23,24,28]. The fourth descriptor is the configurational entropy of mixing of the perovskites, used to differentiate the chemical complexity of solid solutions. Tens of compositions were

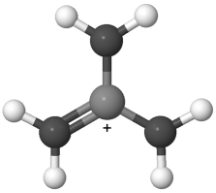
synthesized, and their structures, optoelectronic, and electrical behavior were characterized. The compositions converted to their descriptors were correlated with the measured properties by simple statistical analysis, leading to new insights into the descriptor-property relations in 3D halide perovskites.

2. MATERIALS AND METHODS

For the present study, we synthesized a total of 65 compositions of the type $A_\alpha PbI_3$ with the A-site cations shown in Table 1, such that $A_\alpha = MA_aTMS_bFA_cEA_dAC_eGA_f$ where $\sum_{\alpha\,=\,a}^{f} \alpha = 1$. The structures of all samples were characterized by X-ray diffraction. The details of the synthesis procedure, the composition table, and the respective diffractogram patterns are shown in Supplementary Note 1.

Table 1 - Characteristics of the A-site organic cations used in the compositions. r_A, μ_A, and NH_A are, respectively, the radius, dipole moment, and number of N-H bonds. All the radius values were taken from [14,15], except for TMS^+, which was taken from [39]. Dipole moments were estimated using MolCalc [40]. Other dipole moments can be found in [15].

Name (symbol)	Cation structure	r_A (pm)	μ_A (D)	NH_A
Methylammonium (MA^+)		217	2.69	3
Trimethylsulfonium (TMS^+)		244	1.07	0
Formamidinium (FA^+)		253	0.22	4
Ethylammonium (EA^+)		274	4.97	3
Acetamidinium (AC^+)		277	2.01	4

Guanidinium (GA^+)		278	0	6

Among the 65 investigated compositions, 33 were identified as single-phase materials. These were characterized by differential scanning calorimetry (DSC) to determine the tetragonal-cubic phase transition temperatures, by UV-Vis diffuse reflectance spectroscopy to evaluate the optical band gaps and Urbach energies, and by current-voltage (I-V) electrometry to assess electrical conductivity and hysteresis behavior. Details of the experimental procedures and corresponding results are provided in Supplementary Note 2. To correlate which and how characteristics of the A-site cations affect these five properties of the perovskites formed, the cation formulations were parameterized into four descriptors through the relations $r_{eff} = \sum_{\alpha\,=\,a}^{f} \alpha r_{\mathrm{A}}$, $\mu_{eff} = \sum_{\alpha\,=\,a}^{f} \alpha \mu_{\mathrm{A}}$, $\mathrm{NH}_{eff} = \sum_{\alpha\,=\,a}^{f} \alpha \mathrm{NH}_{\mathrm{A}}$, and $\Delta S_c = -(R/5) \sum_{\alpha\,=\,a}^{f} \alpha(\ln\alpha)$, such that r_{A} and r_{eff} are the individual and effective radius, μ_{eff} is the dipole moment, NH_{eff} and NH_{A} are the effective and individual number of N-H bonds, and ΔS_c is the configurational entropy of mixing, with R the gas constant. The table containing the compositions converted to their descriptors and the respective experimentally determined properties is given in Supplementary Note 3. We stress that, while several other properties could have been studied, these five chosen are among the most important ones for HP materials and devices and are general enough to represent a broad spectrum of their characteristics.

For the sake of completeness, we point out that several other interesting descriptors could be proposed, such as the molecular rigidity [41,42] and the globularity [43], but we decided to use only the above-mentioned descriptors since they are simple and already available in the literature or can be easily obtained for new molecules. Other descriptors, such as the oxidation state of cations, were not considered because it does not differentiate the used species. The oxidation state, as used by Barthel [44], would be important for diammonium cations, such as ethylenediammonium [45,46]. Obviously, varying B and X sites in the ABX_3 stoichiometry would require additional descriptors, such as their sizes, charges, and electronegativities, which would increase the complexity beyond what is currently intended.

Each one of the five properties, designated as P_n, are correlated with the four descriptors through a multivariate linear regression, via partial least squares, in the form of $P_n = k_n + a_n r_{eff} + b_n \mu_{eff} + c_n \mathrm{NH}_{eff} + d_n \Delta S_c$, where k_n, a_n, b_n, c_n, and d_n are the coefficients of the

polynomials. A detailed description of the statistics used is provided in Supplementary Note 4. The final results were derived from 25 compositions remaining after exclusion of measurements in eight of the 33 compositions for each property, as also detailed in the Supplementary Note 4. An important feature of the approach is the correlation of the signals and normalized modulus of the coefficients with their effects on the properties studied. In other words, if for one property a_n is negative, hence the increase in the associated descriptor, i.e., the radius (r_{eff}), decreases the property's value. Analogously, if b_n is close to zero, then the dipole moment (μ_{eff}) effect on the property is negligible. This attribution is possible based on $\mathrm{lm.beta}$ values (Supplementary Note 5), the standardized regression coefficients computed from deviations or similar measures of the variables used. It gives the weighted effects of the independent variable on the dependent variable. Finally, we point out that we chose to use a purely statistical approach to capture direct physicochemical correlations that may not be possible with machine learning approaches depending on method employed. Also, for linear relations, the statistical approach is completely capable of giving the outcomes we initially wanted. Still, we plan to employ machine learning methods in the future to address some specific composition-property relations.

3. RESULTS AND DISCUSSION

Before delving into the quantitative regressions of the properties, some observations on the formability of single-phased perovskites must be made. All compositions containing TMS^+ and the ones with 10% or more of AC^+ content are multiphase. Also, some compositions with AC^+ contents as low as 5% are multiphase when associated with other cations. We previously [47] discussed the formation of solid solutions and solubility limits of mixed A-site compositions based on $MAPbI_3$, where we highlighted that both the radius and number of N-H bonds of the substituting cations are important to obtain single-phased systems. For example, TMS^+ does not form hydrogen bonds with the inorganic framework and, because of that, its solubility in $MAPbI_3$ is essentially negligible. On the other hand, AC^+ has four N-H bonds, but it is too large, hence its solubility limit is only ~10% [48]. Inspired by the concept of the "diverse-ion effect" (or uncommon ion effect), where the solubility of one chemical species in a given solution is increased when different ions are present, to the present paper, we tried to test if we could enhance the solubility limits by increasing the overall entropy of mixing, i.e., by mixing TMS^+ and AC^+ with other cations simultaneously. Moreover, substituting MA^+ with both TMS^+ and GA^+ in principle keeps the system's average number of N-H bonds the same, so

that GA^+ would offset the reduction caused by TMS^+, and maybe we could dissolve a bit more of TMS^+ in the host perovskite. While the idea might be appealing, it apparently doesn't happen in halide perovskites, at least in the $MAPbI_3$-based ones. Regardless of the cations used and their contents, their solubility limits seem to remain unchanged. Even though this is a bit disappointing for now, we will insist on this topic in future work.

The first correlation to be established with the A-site descriptors is the phase transition temperature, in particular, between the tetragonal and cubic phases. It is well known that $MAPbI_3$-based compositions exhibit a tetragonal-to-cubic (T → C) phase transition near room temperature, typically in the range of 250-330 K depending on the substituting cations used and their contents [15]. Other phases may appear, but only at much lower temperatures, around 150 K. Hence, we focus on the phase transitions around room temperature. Beyond its fundamental relevance, tracking stable phases and transitions is important since there is a general preference for cubic halide perovskites for solar cell applications [49]. Moreover, structural and dimensional distortions due to phase transitions at operational temperatures may cause device malfunction or failure. In Figure 1a, we show the correlation between experimental and calculated T → C transition temperatures derived from the proposed descriptors, indicating a clear correspondence, meaning that the A-site characteristics can well describe the transition temperatures. From the perspective of the results shown in Figure 1b, all descriptors influence the T → C temperatures, with N-H bonds being particularly relevant, increasing the temperatures, and entropy of mixing decreasing them. Notably, while increasing the A-site radius reduces the T → C temperatures, as expected from the concept of lattice symmetrization [50], its effect is not as strong as for the mentioned former. Indeed, we showed in a recent paper that AC^+, EA^+, and GA^+ cations, with almost equal radii, have remarkably different effects on the phase transitions, which were already linked to the number of N-H bonds [35]. The proposed explanation was based on increased interactions with the inorganic framework, hindering cation dynamics and the maintenance of the tetragonal distortion, thereby requiring higher temperatures to transition to the cubic phase. This interpretation is consistent with the interplay between size effects and enhanced hydrogen bonding in stabilizing the cubic phase when comparing MA^+ and FA^+ [29]. The behavior with entropy of mixing might be due to the cubic phase being the most stable at higher temperatures, which is favored by a high entropy system.

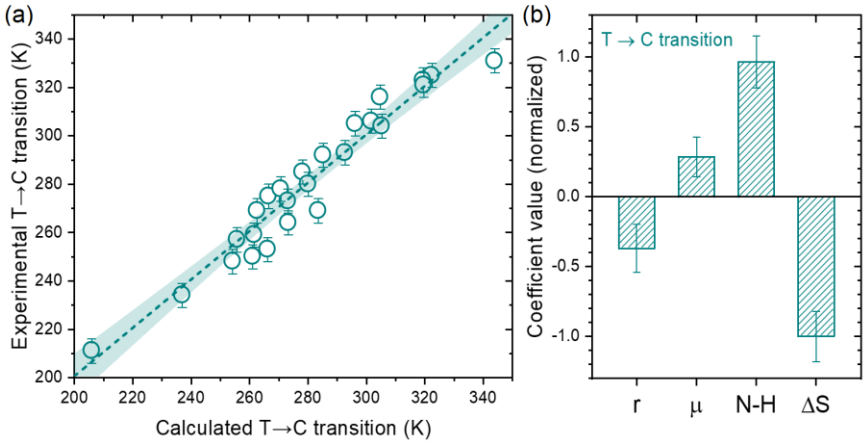


Figure 1 - Phase transitions dependency on the proposed descriptors. a) Experimental versus calculated temperatures of the tetragonal-cubic phase transition. The dashed line is the linear fit (slope = 1), and the shaded area is the 95% confidence level bands. Error bars are experimental uncertainties. b) Normalized coefficient values for each descriptor of the multivariate linear regression polynomial obtained from the experimental tetragonal-cubic phase transition temperatures.

Understanding the optoelectronic properties of halide perovskites is of prominent importance due to their application in electronic devices. In particular, band gap and Urbach energies are among the most important properties in this regard [51,52]. In halide perovskites, band structure is formed mainly by 5p orbital of I and 6s and 6p orbitals of Pb [53]. Because of that, A-site cations have smaller indirect effects. Still, fine-tuning can be achieved by changing the cation composition, as typically reported in literature. Figure 2 shows the correlation between the experimental and calculated optoelectronic parameters and the effects of the descriptors.

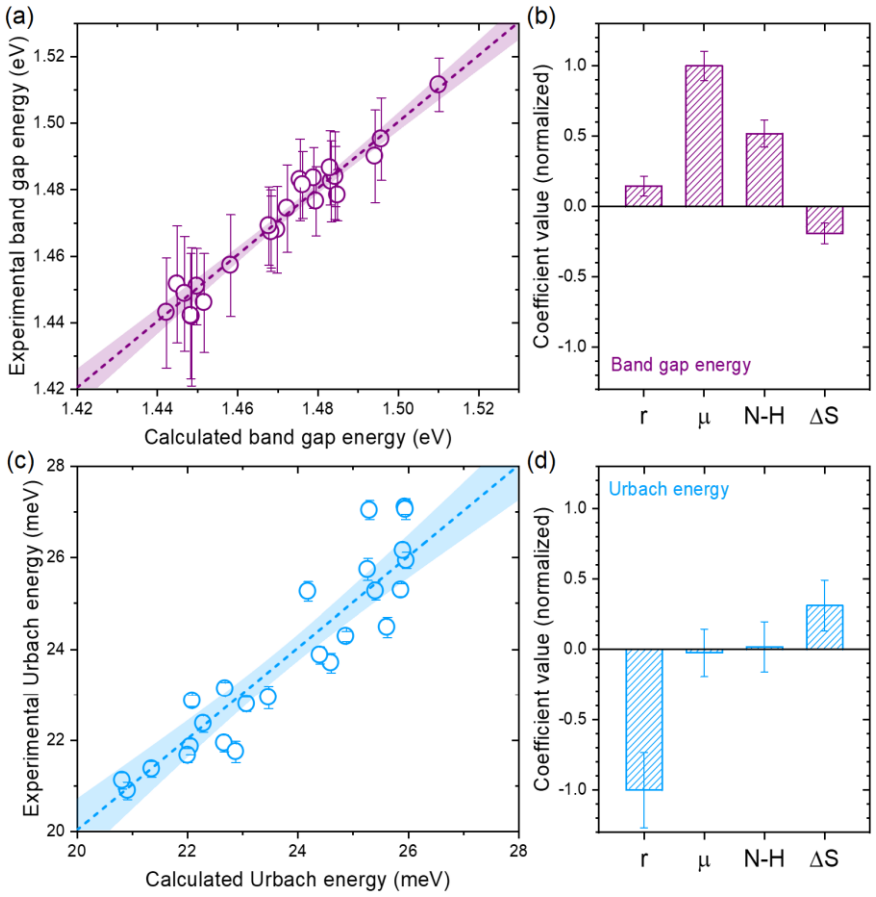


Figure 2 - Optoelectronic properties dependency on the proposed descriptors. Experimental versus calculated a) band gap and c) Urbach energies. Normalized coefficient values for each descriptor of the multivariate linear regression polynomial obtained from the respective experimental b) band gap and d) Urbach energies. The dashed lines are the linear fits (slope = 1), and the shaded areas are the 95% confidence level bands. Error bars are fit residuals.

For the band gap energy (Figures 2a and 2b), it seems that the main controlling factors are the cation dipole and the number of N-H bonds, both of which increase the energy. A possible explanation for the μ_{eff} perspective could be the attraction and repulsion of orbitals by the dipole's electric field, and, in a way resembling the crystal-field theory, this interaction shifts the valence and conduction bands energetically farther apart. For N-H bonds, there may be a connection to the structural distortion imposed by hydrogen bonds, as mentioned earlier, which reduces orbital overlap. Regardless of the reasons, the results may help explain why substituting MA^+ by FA^+ reduces the band gap energy, whereas substituting GA^+ increases it [31]. Both substituting cations have much lower dipole moments than MA^+, but GA^+ has more N-H bonds than FA^+. In this sense, the expected reduction due to lower μ_{eff} is observed for FA^+, but not for GA^+. Interestingly, the band gap energy in $EA_xMA_{1-x}PbBr_3$ nanocrystals increases sensibly, from 2.38 to 2.94 eV, with the EA^+ content, which was attributed to EA^+

being larger than MA^+ [54]. From our data, the dominant reason for this behavior may be the higher dipole of EA^+. In any case, studies on band gap engineering of HPs are easily found in the literature [55-59]. On the other hand, this approach applied specifically to mixed A-site HPs is rather scarce. We plan to dig deeper into this topic in the near future to refine our insights into descriptor-property relations.

For the Urbach energy (Figures 2c and 2d), the effects are intriguingly the opposite, with radius and the entropy of mixing controlling the property. The Urbach energy, a contribution from sub-band gap (band-tailing) absorption, can be viewed as a metric for lattice disorder and can result from thermal disorder or from structural imperfections [60]. From this perspective, structural heterogeneity must play a significant role. Hence, the effect of controlling descriptors can be viewed as a consequence of increasing the radius, leading to lattice symmetrization, and of entropy of mixing, leading to local non-homogeneity. In the $Cs_xFA_yMA_{1-x-y}$ system, it was found that FA^+ decreases the Urbach energy, while Cs^+ increases it [32], which agrees with the proposed descriptors. In the $GA_xFA_xMA_{1-2x}$ system, x values up to ~ 0.08 reduces the Urbach energy, but after it starts to increase [61]. From the perspective of the present results, it might be explained by entropy effects dominating at higher substitution levels after the earlier prevalence of larger cations.

Lastly, we studied charge-transport properties based on analyses of electrical conductivity and I-V hysteresis curves. The results are summarized in Figure 3. Unlike other properties discussed so far, charge transport measurements and results are very sensitive to several internal (microstructure, defects, etc.) and external (composition and geometry of electrodes, scanning rates, etc.) factors, none of which are included in our models. Despite our best efforts to obtain representative results, understanding and modeling the data are quite challenging. As such, the data variance is way larger than in the other cases. Still, it doesn't mean we cannot draw some thoughts on this matter.

Although electrical conductivity and hysteresis are largely due to different charge carriers, electronic and ionic A-site cations should affect both similarly, as shown previously [31,33-35]. Hence, it is expected that both electrical conductivity (Figures 3a and 3b) and hysteresis (Figures 3c and 3d) seem to be controlled essentially by the same descriptors, namely the dipole and number of N-H bonds. A deviation occurs in the radius effect, which we also expected to reduce hysteresis. Most intriguingly, the entropy of mixing apparently plays little role in either charge-transport characteristic.

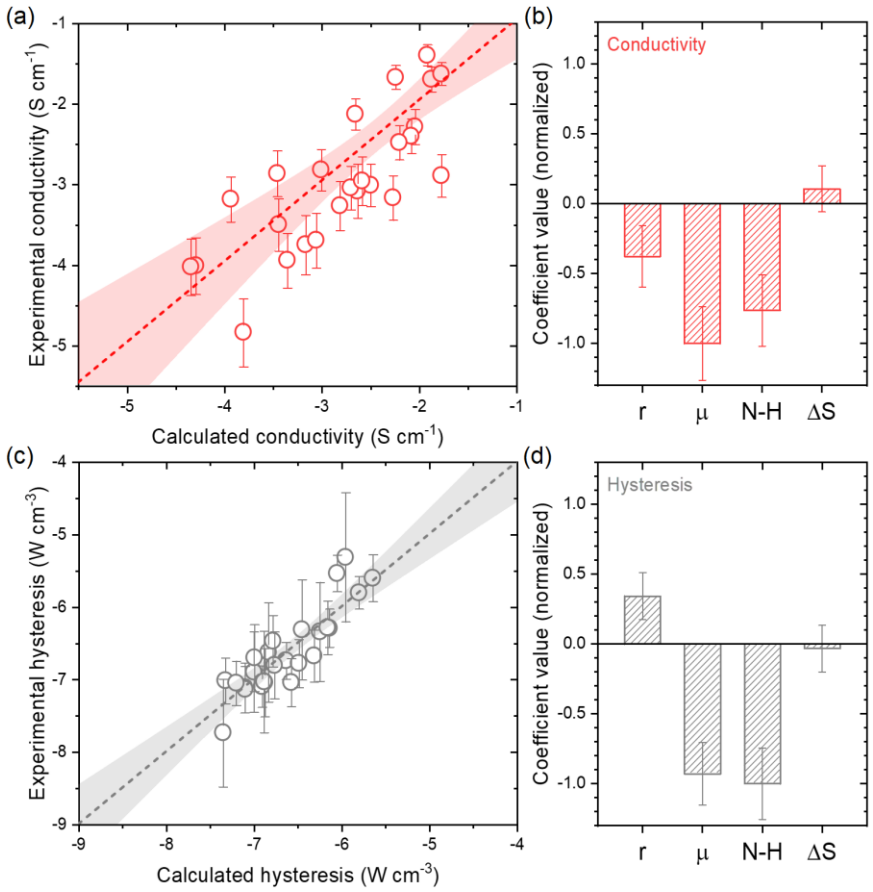


Figure 3 - Charge transport properties dependency on the proposed descriptors. Experimental versus calculated a) electrical conductivity and c) I-V hysteresis. Values are on a logarithmic scale. Normalized coefficient values for each descriptor of the multivariate linear regression polynomial obtained from the respective experimental b) electrical conductivity and d) I-V hysteresis. The dashed lines are the linear fits (slope = 1), and the shaded areas are the 95% confidence level bands. Error bars are standard deviations.

Carrier scattering is known to be dominated by discontinuities, which are inherently connected to configurational entropy. A closer inspection of our data, as given in Supplementary Note 5, shows that the entropy of mixing was relevant in the raw data, i.e., before data selection. It is possible that, in this case, the outcome may have been a bit perturbed. On the other hand, we recently reported results on halide perovskite compositions with maximized entropy, where electrical conductivity and hysteresis obtained through equivalent measurements revealed that the observed changes are weaker with increased entropy than in systems of lower entropies [62]. By comparison, AC^+ alone appears to be more effective at reducing both electrical conductivity and hysteresis than simultaneous four-cation substitution with similar total substitution content [35,48]. In this sense, it may be possible that the entropy effect is indeed minor. In conclusion, it will remain until more exploration on the topic.

Following, similar results regarding the dipole effect were reported earlier [35] and were attributed to interactions between charge carriers and the electric fields of the dipoles (similar to a dielectric drag effect [63]), reducing carrier mobility. For N-H bonds, the effect on the electrical conductivity may be due to the tetragonal distortion. As shown in Supplementary Note 6, on average, compositions in the tetragonal phase (lower symmetry) at room temperature exhibit lower electrical conductivity than those in the cubic ones (higher symmetry). Peculiarly, we found that compositions of cubic phases at room temperature tend to exhibit lower hysteresis, band gaps, and Urbach energies, corroborating the general desirability of cubic perovskites for solar cell applications. Lastly, the effects of N-H bonds in hysteresis can be attributed to more interactions with the I^- ions in the inorganic framework. Since mobile iodide ions and vacancies are the main ionic migrating species [64], forces that keep I^- ions in their positions will mitigate overall ion transport, and therefore hysteresis.

Having all these scenarios set, a summary of the main findings relating the studied properties to the A-site descriptors is given in Table 2. We must point out that these relations should be considered having in mind that they were derived from statistical analyses instead of from mechanistic studies. From the material's perspective, some possibly significant limitations are disregarding double bonds as descriptors and assuming that all N-H bonds are equal and, implicitly, that they can form hydrogen bonds evenly, thereby neglecting spatial (steric, geometric, orientational, etc.) requirements. Moreover, the ionic radii of organic cations are not strictly defined in the sense that they may very well be a distribution of values [65]. Also, since molecules are not spherically symmetric, the assumed ionic radii implicitly consider their isotropic rotational dynamics. Finally, the effective A-site characteristics were assumed to be simple weighted averages (Vegard's Law) of those of the cations used. Although this is commonly done for radii, there is no guarantee that this is the best transformation method for dipoles and the number of N-H bonds. From a statistical/modeling perspective, the descriptors should ideally be completely independent variables, but there is some inherent collinearity because we cannot change the parameters completely freely. To have "less dependent" variables, additional cations would preferentially be used. Additionally, most other cations would not form single-phase perovskites, which is a requisite for the present type of study. Also, an obvious drawback is the small data set, which forced us to use z-score normalization to obtain better R^2 values. By doing so, quantitative predictions of additional compositions would require including the new data in the original data set, thereby changing the z-scored values and, in turn, the coefficients of the linear model. Hence, a predictive approach for additional

compositions can only be qualitative. Optimally, the larger the data set, the better the results would be and a quantitative outcome could be obtained. Finally, we assumed a linear relation between descriptors and properties, but non-linear relations could also have been used. On the other hand, this approach would increase the complexity beyond the intended purposes of this work.

Table 2 - Summary of descriptor effects on the studied properties of halide perovskites. Two arrows indicate a strong effect (normalized coefficient > 0.5), one arrow for a weak effect (normalized coefficient between 0.2 and 0.5), and the signal (-) means that the effect is negligible (normalized coefficients < 0.2). Up arrow means that the increase of the descriptor increases the property value, and the down arrow means it reduces the property value. The symbol (*) denotes inconclusive results, as discussed in the main text.

Property	**What does the increase of the A-site descriptor cause?**			
	r_{eff}	μ_{eff}	NH_{eff}	ΔS_c
T → C transition	↓	↑	↑↑	↓↓
Band gap energy	↑	↑↑	↑↑	↓
Urbach energy	↓↓	-	-	↑
Electrical conductivity	↓	↓↓	↓↓	*
Hysteresis	↑	↓↓	↓↓	*

It is important to note that the present study used only organic cations, all of which are inherently anisotropic. As a result, the insights may not be adequately applied to systems containing isotropic inorganic cations. Hence, predictions based solely on organic cations should be preferred. With due consideration, an attempt to make some predictions about hypothetical systems composed of the heteroatom-containing cations fluoromethylammonium (FMA) and hydroxymethylammonium (HMA) is provided in Supplementary Note 7.

It should be noted that, although we used lead iodide-based compositions, we expect the same principles and rationale can be applied when changing the A-site in other systems, such as those containing Sn^{2+} at the B-site and Br^- and Cl^- at the X-site, differing only in absolute terms. One example, as discussed in Supplementary Note 8, is the use of DMA^+ and FA^+ as substituents in $MAPbBr_3$-based perovskites leads to variations within the predictions that can be made with the relationships found here. Even so, the exploration of these possibilities should be the subject of future research, allowing the inclusion of other descriptors such as the radii and electronegativities of the B and X components. Noteworthy, the correlations searched in the present work cannot be applied to other classes of perovskite-like materials, such as double

and LD perovskites, which are structurally more complex and would demand specific works to do so.

4. CONCLUSIONS

In summary, using a combination of synthesis and experimental characterization of dozens of $MAPbI_3$-based compositions and statistical tools, we have shown how the structural, optoelectronic, and charge-transport properties of halide perovskites can be described and controlled in terms of the A-site cation characteristic. Some relevant effects found were the entropy in reducing and number of N-H bonds in increasing the phase transition temperature, the dipole moment in increasing the band gap, the radius in reducing and the entropy in increasing the Urbach energy and the dipole moment and N-H bonds in reducing the conductivity and I-V hysteresis. While we made significant progress in understanding A-site composition-property relations in 3D ABX_3 perovskites, our main intention is to stimulate the specialized research community to follow a similar path to elucidate these links. Studying other properties and using additional compositions and descriptors can be readily achieved through automated laboratories and/or high-throughput calculations, yielding even better results needed to optimize the composition engineering of perovskite materials and devices.

Author contributions

F. B. Minussi: conceptualization (leading), methodology (leading), validation (leading), formal analysis (leading), investigation (supporting), writing - original draft (leading), writing - review and editing (leading), and visualization (leading). R. M. Silva Jr.: conceptualization (support), methodology (support), software (leading), validation (support), formal analysis (leading), investigation (leading), writing - original draft (support), writing - review and editing (support), and visualization (support). M. A.M. Teixeira: investigation (support), writing - original draft (support), writing - review and editing (support). R. P. Bonini: investigation (support), writing - original draft (support), writing - review and editing (support). T. A. Abreu: conceptualization (support), methodology (support), software (leading), validation (support), formal analysis (support), writing - original draft (support), writing - review and editing (support). J. A. Eiras: conceptualization (support), methodology (support), resources (leading), writing - original draft (support), writing - review and editing (support), supervision (leading), project administration (leading), and funding acquisition (leading). E. B. Araújo: conceptualization (support), methodology (support), resources (leading), writing - original draft (support), writing - review

and editing (support), supervision (leading), project administration (leading), and funding acquisition (leading).

Data availability

The data supporting the findings of this study are available within the article and its electronic Supplementary Information.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work has been supported by Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-INCT: 406322/2022-8 and CNPq Grant: 315586/2023-0) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-DAAD: 88881.895061/2023-01). The author F.B. Minussi would like to thank the CNPq for their Postdoctoral Fellowship (Grant: 157815/2025-0). The authors thank Mr. Elton José de Souza for supporting the SEM and DSC experiments.

5. REFERENCES

[1] A. K. Jena, A. Kulkarni and T. Miyasaka, Halide Perovskite Photovoltaics: Background, Status, and Future Prospects, *Chem. Rev.*, 2019, **119**, 3036–3103.

[2] X.-K. Liu, W. Xu, S. Bai, Y. Jin, J. Wang, R. H. Friend and F. Gao, Metal halide perovskites for light-emitting diodes, *Nat. Mater.*, 2021, **20**, 10–21.

[3] X. He, Y. Deng, D. Ouyang, N. Zhang, J. Wang, A. A. Murthy, I. Spanopoulos, S. M. Islam, Q. Tu, G. Xing, Y. Li, V. P. Dravid and T. Zhai, Recent Development of Halide Perovskite Materials and Devices for Ionizing Radiation Detection, *Chem. Rev.*, 2023, **123**, 1207–1261.

[4] S.-Y. Kim, H. Zhang and J. Rubio-Magnieto, Operating Mechanism Principles and Advancements for Halide Perovskite-Based Memristors and Neuromorphic Devices, *J. Phys. Chem. Lett.*, 2024, **15**, 10087–10103.

[5] N. Tewari, D. Lam, C. H. A. Li and J. E. Halpert, Recent advancements in batteries and photo-batteries using metal halide perovskites, *APL Materials*, 2022, **10**, 040905.

[6] B. Li, L. Qiao, H. Yang, Y. Wang, W. Gu and B. Du, Perovskite materials empower sensors and batteries: environmental health monitoring and flexible wearable innovation, *J. Mater. Chem. C*, 2025, **13**, 14727–14742.

[7] Y. Chen, Z. Chen, X. Zhang, J. Chen and Y. Wang, An organic-halide perovskite-based photo-assisted Li-ion battery for photoelectrochemical storage, *Nanoscale*, 2022, **14**, 10903–10909.

[8] P. Hänsch and M. A. Loi, Metal halide perovskites: A rising platform for gas sensing?, *Applied Physics Letters*, 2023, **123**, 030501.

[9] P. Singh, N. K. Bansal, S. Dey, R. Singh and T. Singh, Recent Progress on Perovskite Materials for VOC Gas Sensing, *Langmuir*, 2024, **40**, 21931–21956.

[10] X. Wang, Y. Peng, S. Yang, H. G. Yang and Y. Hou, Recent progress in metal halide perovskite photocatalysts for hydrogen evolution, *Mater. Chem. Front.*, 2023, **7**, 4635–4657.

[11] M. A. Raza, F. Li, M. Que, L. Zhu and X. Chen, Photocatalytic reduction of CO2 by halide perovskites: recent advances and future perspectives, *Mater. Adv.*, 2021, **2**, 7187–7209.

[12] J. Breternitz and S. Schorr, What Defines a Perovskite?, *Advanced Energy Materials*, 2018, **8**, 1802366.

[13] Q. A. Akkerman and L. Manna, What Defines a Halide Perovskite?, *ACS Energy Lett.*, 2020, **5**, 604–610.

[14] G. Kieslich, S. Sun and A. K. Cheetham, Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog, *Chem. Sci.*, 2014, **5**, 4712–4715.

[15] M. Simenas, A. Gagor, J. Banys and M. Maczka, Phase Transitions and Dynamics in Mixed Three- and Low-Dimensional Lead Halide Perovskites, *Chem. Rev.*, 2024, **124**, 2281–2326.

[16] P. Liu, X. Li, T. Cai, W. Xing, N. Yang, H. Arandiyan, Z. Shao, S. Wang and S. Liu, Molecular Structure Tailoring of Organic Spacers for High-Performance Ruddlesden–Popper Perovskite Solar Cells, *Nano-Micro Lett.*, 2025, **17**, 35.

[17] C. Wang, X. Dong, F. Chen, G. Liu and H. Zheng, Recent progress of two-dimensional Ruddlesden–Popper perovskites in solar cells, *Mater. Chem. Front.*, 2023, **7**, 5786–5805.

[18] L. Mao, C. C. Stoumpos and M. G. Kanatzidis, Two-Dimensional Hybrid Halide Perovskites: Principles and Promises, *J. Am. Chem. Soc.*, 2019, **141**, 1171–1190.

[19] H. Fu, Dion–Jacobson halide perovskites for photovoltaic and photodetection applications, *J. Mater. Chem. C*, 2021, **9**, 6378–6394.

[20] F. Ünlü, E. Jung, J. Haddad, A. Kulkarni, S. Öz, H. Choi, T. Fischer, S. Chakraborty, T. Kirchartz and S. Mathur, Understanding the interplay of stability and efficiency in A-site engineered lead halide perovskites, *APL Materials*, 2020, **8**, 070901.

[21] J. Gu and Y. Fu, Is There an Optimal Spacer Cation for Two-Dimensional Lead Iodide Perovskites?, *ACS Mater. Au*, 2025, **5**, 24–34.

[22] X. Li, J. M. Hoffman and M. G. Kanatzidis, The 2D Halide Perovskite Rulebook: How the Spacer Influences Everything from the Structure to Optoelectronic Device Efficiency, *Chem. Rev.*, 2021, **121**, 2230–2291.

[23] J.-W. Lee, S. Tan, S. I. Seok, Y. Yang and N.-G. Park, Rethinking the A cation in halide perovskites, *Science*, 2022, **375**, eabj1186.

[24] D. A. Egger, A. Bera, D. Cahen, G. Hodes, T. Kirchartz, L. Kronik, R. Lovrincic, A. M. Rappe, D. R. Reichman and O. Yaffe, What Remains Unexplained about the Properties of Halide Perovskites?, *Advanced Materials*, 2018, **30**, 1800691.

[25] M. M. Byranvand, C. Otero-Martínez, J. Ye, W. Zuo, L. Manna, M. Saliba, R. L. Z. Hoye and L. Polavarapu, Recent Progress in Mixed A-Site Cation Halide Perovskite Thin-Films and Nanocrystals for Solar Cells and Light-Emitting Diodes, *Advanced Optical Materials*, 2022, **10**, 2200423.

[26] Y. Ma, Y. Wang, Y. Fang, Y. Jiang, Z. Dai and J. Miao, Recent Advances in the A-Site Cation Engineering of Lead Halide Perovskites, *Advanced Optical Materials*, 2024, **12**, 2401367.

[27] M. P. Hautzinger, W. Mihalyi-Koch and S. Jin, A-Site Cation Chemistry in Halide Perovskites, *Chem. Mater.*, 2024, **36**, 10408–10420.

[28] O. I. Kucheriv, V. Y. Sirenko and I. A. Gural'skiy, Small Cations for Engineering 3D Hybrid Perovskites, *Chemistry A European J*, 2025, **31**, e202500765.

[29] A. Amat, E. Mosconi, E. Ronca, C. Quarti, P. Umari, Md. K. Nazeeruddin, M. Grätzel and F. De Angelis, Cation-Induced Band-Gap Tuning in Organohalide Perovskites: Interplay of Spin–Orbit Coupling and Octahedra Tilting, *Nano Lett.*, 2014, **14**, 3608–3616.

[30] J. Han, K. Park, S. Tan, Y. Vaynzof, J. Xue, E. W.-G. Diau, M. G. Bawendi, J.-W. Lee and I. Jeon, Perovskite solar cells, *Nat Rev Methods Primers*, 2025, **5**, 1–27.

[31] F. B. Minussi, R. M. Silva and E. B. Araújo, Interplay between cation composition and charge transport characteristics in GA *x* FA *y* MA1- *x* - *y* PbI3 halide perovskites, *Applied Physics Letters*, 2024, **124**, 153901.

[32] S. Moradi, S. Kundu, M. Awais, Y. Haruta, H. Nguyen, D. Zhang, F. Tan and M. I. Saidaminov, High-Throughput Exploration of Triple-Cation Perovskites via All-in-One Compositionally-Graded Films, *Small*, 2023, **19**, 2301037.

[33] F. B. Minussi, R. M. Da Silva and E. B. Araújo, Differing Effects of Mixed A-Site Composition on the Properties of Hybrid Lead Iodide Perovskites, *J. Phys. Chem. C*, 2023, **127**, 8814–8824.

[34] F. B. Minussi, R. M. Silva and E. B. Araújo, Composition-Property Relations for $GA_x FA_y MA_{1-x-y} PbI_3$ Perovskites, *Small*, 2024, **20**, 2305054.

[35] F. B. Minussi, R. M. Silva, J. C. S. Moraes and E. B. Araújo, Organic cations in halide perovskite solid solutions: exploring beyond size effects, *Phys. Chem. Chem. Phys.*, 2024, **26**, 20770–20784.

[36] S. Jin, Can We Find the Perfect A-Cations for Halide Perovskites?, *ACS Energy Lett.*, 2021, **6**, 3386–3389.

[37] J. Ibaceta-Jaña, M. Chugh, A. S. Novikov, H. Mirhosseini, T. D. Kühne, B. Szyszka, M. R. Wagner and R. Muydinov, Do Lead Halide Hybrid Perovskites Have Hydrogen Bonds?, *J. Phys. Chem. C*, 2022, **126**, 16215–16226.

[38] L. Lin, T. W. Jones, T. C.-J. Yang, X. Li, C. Wu, Z. Xiao, H. Li, J. Li, J. Qian, L. Lin, J. Q. Shi, S. D. Stranks, G. J. Wilson and X. Wang, Hydrogen bonding in perovskite solar cells, *Matter*, 2024, **7**, 38–58.

[39] A. Kaltzoglou, C. C. Stoumpos, A. G. Kontos, G. K. Manolis, K. Papadopoulos, K. G. Papadokostaki, V. Psycharis, C. C. Tang, Y.-K. Jung, A. Walsh, M. G. Kanatzidis and P. Falaras, Trimethylsulfonium Lead Triiodide: An Air-Stable Hybrid Halide Perovskite, *Inorg. Chem.*, 2017, **56**, 6302–6309.

[40] J. H. Jensen and J. C. Kromann, The Molecule Calculator: A Web Application for Fast Quantum Mechanics-Based Estimation of Molecular Properties, *J. Chem. Educ.*, 2013, **90**, 1093–1095.

[41] M. B. Fridriksson, N. Van Der Meer, J. De Haas and F. C. Grozema, Tuning the Structural Rigidity of Two-Dimensional Ruddlesden–Popper Perovskites through the Organic Cation, *J. Phys. Chem. C*, 2020, **124**, 28201–28209.

[42] L. Scalon, C. A. Nogueira, A. F. V. Fonseca, P. E. Marchezi, R. F. Moral, G. Grancini, T. Kodalle, C. M. Sutter-Fella, C. C. Oliveira, L. F. Zagonel and A. F. Nogueira, 2D Phase Formation on 3D Perovskite: Insights from Molecular Stiffness, *ACS Appl. Mater. Interfaces*, 2024, **16**, 51727–51737.

[43] S. Gholipour, A. M. Ali, J. Correa-Baena, S. Turren-Cruz, F. Tajabadi, W. Tress, N. Taghavinia, M. Grätzel, A. Abate, F. De Angelis, C. A. Gaggioli, E. Mosconi, A. Hagfeldt and M. Saliba, Globularity-Selected Large Molecules for a New Generation of Multication Perovskites, *Advanced Materials*, 2017, **29**, 1702005.

[44] C. J. Bartel, C. Sutton, B. R. Goldsmith, R. Ouyang, C. B. Musgrave, L. M. Ghiringhelli and M. Scheffler, New tolerance factor to predict the stability of perovskite oxides and halides, *Sci. Adv.*, 2019, **5**, eaav0693.

[45] G. Rimkus, S. Balčiūnas, M. Ptak, S. Smółka, V. Klimavičius, D. Szewczyk, A. Sieradzki, V. Kalendra, J. Banys, M. Mączka and M. Šimėnas, Phase Diagram and Dielectric Response of Hybrid Lead Halide Hollow Perovskites: A Universal Behavior of Molecular Cation Mixing, *Chem. Mater.*, 2024, **36**, 7397–7405.

[46] I. Spanopoulos, W. Ke, C. C. Stoumpos, E. C. Schueller, O. Y. Kontsevoi, R. Seshadri and M. G. Kanatzidis, Unraveling the Chemical Nature of the 3D "Hollow" Hybrid Halide Perovskites, *J. Am. Chem. Soc.*, 2018, **140**, 5728–5742.

[47] F. Brondani Minussi, R. M. da Silva and E. Borges Araújo, Does it mix? Insights and attempts to predict the formability of single phase mixed A-cation lead iodide perovskites, *Inorganic Chemistry Frontiers*, 2023, **10**, 6129–6133.

[48] F. B. Minussi, R. C. D. O. Santos, M. A. D. M. Teixeira, R. M. Silva and E. B. Araújo, An Exploration of the Role of Acetamidinium Substitution in Methylammonium Lead Iodide Perovskites, *ChemPhysChem*, 2026, **27**, e202500259.

[49] Q. Wei, H. Liang, Y. Haruta, M. Saidaminov, Q. Mi, M. Saliba, G. Cui and Z. Ning, From tetragonal to cubic: perovskite phase structure evolution for high-performance solar cells, *Science Bulletin*, 2023, **68**, 141–145.

[50] M. Mączka, M. Ptak, K. Fedoruk, D. Stefańska, A. Gągor, J. K. Zaręba and A. Sieradzki, The lattice symmetrization worked, but with a plot twist: effects of methylhydrazinium doping of MAPbI3 on phase transitions, cation dynamics and photoluminescence, *J. Mater. Chem. C*, 2024, **12**, 1396–1405.

[51] O. Almora, G. C. Bazan, C. I. Cabrera, L. A. Castriotta, S. Erten-Ela, K. Forberich, K. Fukuda, F. Guo, J. Hauch, A. W. Y. Ho-Baillie, T. J. Jacobsson, R. A. J. Janssen, T. Kirchartz, R. R. Lunt, X. Mathew, D. B. Mitzi, M. K. Nazeeruddin, J. Nelson, A. F. Nogueira, Ulrich. W. Paetzold, B. P. Rand, U. Rau, T. Someya, C. Sprau, L. Vaillant-Roca and C. J. Brabec, Device Performance of Emerging Photovoltaic Materials (Version 5), *Advanced Energy Materials*, 2025, **15**, 2404386.

[52] S. De Wolf, J. Holovsky, S.-J. Moon, P. Löper, B. Niesen, M. Ledinsky, F.-J. Haug, J.-H. Yum and C. Ballif, Organometallic Halide Perovskites: Sharp Optical Absorption Edge and Its Relation to Photovoltaic Performance, *J. Phys. Chem. Lett.*, 2014, **5**, 1035–1039.

[53] A. Walsh, Principles of Chemical Bonding and Band Gap Engineering in Hybrid Organic–Inorganic Halide Perovskites, *J. Phys. Chem. C*, 2015, **119**, 5755–5760.

[54] M. Mittal, A. Jana, S. Sarkar, P. Mahadevan and S. Sapra, Size of the Organic Cation Tunes the Band Gap of Colloidal Organolead Bromide Perovskite Nanocrystals, *J. Phys. Chem. Lett.*, 2016, **7**, 3270–3277.

[55] Z. Hu, Z. Lin, J. Su, J. Zhang, J. Chang and Y. Hao, A Review on Energy Band-Gap Engineering for Perovskite Photovoltaics, *Solar RRL*, 2019, **3**, 1900304.

[56] H. Wang, A. Tal, T. Bischoff, P. Gono and A. Pasquarello, Accurate and efficient band-gap predictions for metal halide perovskites at finite temperature, *npj Comput Mater*, 2022, **8**, 1–13.

[57] F. Gou, Z. Ma, Q. Yang, H. Du, Y. Li, Q. Zhang, W. You, Y. Chen, Z. Du, J. Yang, N. He, J. Luo, Z. Liu, Z. Tian, M. Mao, K. Liu, J. Yu, A. Zhang, F. Min, K. Sun and N. Xuan, Machine Learning-Assisted Prediction and Control of Bandgap for Organic–Inorganic Metal Halide Perovskites, *ACS Appl. Mater. Interfaces*, 2025, **17**, 18383–18393.

[58] R. Zhang, B. Motes, S. Tan, Y. Lu, M.-C. Shih, Y. Hao, K. Yang, S. Srinivasan, M. G. Bawendi and V. Bulović, Machine Learning Prediction of Organic–Inorganic Halide Perovskite Solar Cell Performance from Optical Properties, *ACS Energy Lett.*, 2025, **10**, 1714–1724.

[59] C. Yang, X. Chong, M. Hu, W. Yu, J. He, Y. Zhang, J. Feng, Y. Zhou and L.-W. Wang, Accelerating the Discovery of Hybrid Perovskites with Targeted Band Gaps via Interpretable Machine Learning, *ACS Appl. Mater. Interfaces*, 2023, **15**, 40419–40427.

[60] E. Ugur, M. Ledinský, T. G. Allen, J. Holovský, A. Vlk and S. De Wolf, Life on the Urbach Edge, *J. Phys. Chem. Lett.*, 2022, **13**, 7702–7711.

[61] F. B. Minussi, L. A. Silva and E. B. Araújo, Structure, optoelectronic properties and thermal stability of the triple organic cation GA x FA x MA 1- 2x PbI 3 system prepared by mechanochemical synthesis, *Physical Chemistry Chemical Physics*, 2022, **24**, 4715–4728.

[62] F. B. Minussi and E. B. Araújo, On the prospects of high-entropy organic A-site halide perovskites, *Phys. Chem. Chem. Phys.*, 2024, **26**, 26479–26488.

[63] M. Bonn, K. Miyata, E. Hendry and X.-Y. Zhu, Role of Dielectric Drag in Polaron Mobility in Lead Halide Perovskites, *ACS Energy Lett.*, 2017, **2**, 2555–2562.

[64] C. Eames, J. M. Frost, P. R. F. Barnes, B. C. O'Regan, A. Walsh and M. S. Islam, Ionic transport in hybrid lead iodide perovskite solar cells, *Nat Commun*, 2015, **6**, 7497.

[65] J. Gebhardt and A. M. Rappe, Mix and Match: Organic and Inorganic Ions in the Perovskite Lattice, *Advanced Materials*, 2019, **31**, 1802697.

SUPPORTING INFORMATION

A-site composition-property relations in halide perovskites: which descriptors do what?

Fernando B. Minussi[1,2*#], Rogerio M. Silva Jr.[3#], Marco Antonio M. Teixeira[1], Ricardo P. Bonini[2], Thays A. Abreu[4], Jose A. Eiras[2], Eudes B. Araujo[1]

[1] Department of Physics and Chemistry, São Paulo State University, Ilha Solteira, SP, Brazil

[2] Department of Physics, Federal University of São Carlos, São Carlos, SP, Brazil

[3] Department of Electrical Engineering, São Paulo State University, Ilha Solteira, SP, Brazil

[4] Department of Mathematics, São Paulo State University, Ilha Solteira, SP, Brazil

CONTENTS

Supplementary Note 1: Composition and phase data of the synthesized perovskites _____ 1

Supplementary Note 2: Experimental data and determination of properties _____ 8

Supplementary Note 3: Effective descriptors and experimental properties _____ 22

Supplementary Note 4: Statistical approach _____ 23

Supplementary Note 5: Normalized coefficients and lm. beta values _____ 25

Supplementary Note 6: Properties in tetragonal and cubic phases _____ 27

Supplementary Note 7: Qualitative predictions of hypothetical systems _____ 28

Supplementary Note 8: Comparison of properties of literature _____ 29

References _____ 30

* fbminussi@gmail.com / f.minussi@unesp.br

# Authors contributed equally

Supplementary Note 1: Composition and phase data of the synthesized perovskites

To synthesize the 65 compositions (Table S1), the precursors used were methylammonium iodide (MAI, Sigma-Aldrich, 99%), guanidinium iodide (MAI, Sigma-Aldrich, 99%), formamidinium iodide (FAI, Sigma-Aldrich, 99%), ethylammonium iodide (EAI, Sigma-Aldrich, 98%), acetamidinium iodide (ACI, Sigma-Aldrich, 98%), trimethylsulfonium iodide (TMSI, Sigma-Aldrich, 98%), and lead iodide (PbI_2, Sigma-Aldrich, 99%). Synthesis proceeded using a mechanochemical process, involving mixing a total of about 400 mg of precursors, in the desired proportions, in a polytetrafluoroethylene container with small zirconia spheres, followed by gently stirring for approximately 40 min. This resulted in a black powder, which was then sieved through a 250 μm mesh, totaling about 350 mg for each composition. Subsequently, pellets with a mass of approximately 150 mg and a diameter of 10 mm were formed using a metal mold and compacted at a pressure of 250 MPa. Finally, the pellets were sintered in a furnace with a heating rate of 5 °C min^{-1} up to 120 °C with isotherm for 2 h, followed by a slow cooling to room temperature (R.T., ~ 25 °C).

Table S1 - Nominal A-site composition of synthesized perovskites and preliminary phase information derived com X-ray diffraction data (given in the sequence).

Number	A-site composition						Single phase?	Symmetry at R.T.
	GA^+	EA^+	AC^+	TMS^+	FA^+	MA^+		
1	-	-	-	-	-	1.00	Yes	Tetragonal
2	0.02	-	-	0.02	-	0.96	No	-
3	0.04	-	-	0.04	-	0.92	No	-
4	0.06	-	-	0.06	-	0.88	No	-
5	-	-	-	0.02	0.02	0.96	No	-
6	-	-	-	0.04	0.04	0.92	No	-
7	-	-	-	0.06	0.06	0.88	No	-
8	-	0.02	-	0.02	-	0.96	No	-
9	-	0.04	-	0.04	-	0.92	No	-
10	-	0.06	-	0.06	-	0.88	No	-
11	-	-	0.02	0.02	-	0.96	No	-
12	-	-	0.04	0.04	-	0.92	No	-
13	-	-	0.06	0.06	-	0.88	No	-
14	0.05	-	0.05	-	-	0.90	Yes	Tetragonal
15	0.10	-	0.10	-	-	0.80	No	-
16	0.15	-	0.15	-	-	0.70	No	-
17	-	-	0.05	-	0.05	0.90	Yes	Tetragonal
18	-	-	0.10	-	0.10	0.80	No	-
19	-	-	0.15	-	0.15	0.70	No	-
20	-	0.05	0.05	-	-	0.90	Yes	Tetragonal
21	-	0.10	0.10	-	-	0.80	No	-
22	-	0.15	0.15	-	-	0.70	No	-
23	0.05	0.05	-	-	-	0.90	Yes	Tetragonal
24	0.10	0.10	-	-	-	0.80	Yes	Cubic
25	0.15	0.15	-	-	-	0.70	Yes	Cubic

Table S1 - Continued.

26	0.20	0.20	-	-	-	0.60	No	-
27	-	0.05	-	-	0.05	0.90	Yes	Cubic
28	-	0.10	-	-	0.10	0.80	Yes	Cubic
29	-	0.15	-	-	0.15	0.70	Yes	Cubic
30	-	0.20	-	-	0.20	0.60	Yes	Cubic
31	-	0.25	-	-	0.25	0.50	Yes	Cubic
32	-	-	-	-	0.50	0.50	Yes	Cubic
33	-	-	-	0.03	0.50	0.47	No	-
34	-	-	-	0.06	0.50	0.44	No	-
35	-	-	-	0.10	0.50	0.40	No	-
36	-	-	0.05	-	0.50	0.45	No	-
37	-	-	0.10	-	0.50	0.40	No	-
38	-	-	0.15	-	0.50	0.35	No	-
39	-	0.05	-	-	0.50	0.45	Yes	Cubic
40	-	0.10	-	-	0.50	0.40	Yes	Cubic
41	-	0.15	-	-	0.50	0.35	No	-
42	0.05	-	-	-	0.50	0.45	Yes	Cubic
43	0.10	-	-	-	0.50	0.40	Yes	Cubic
44	0.15	-	-	-	0.50	0.35	Yes	Cubic
45	0.10	-	-	-	-	0.90	Yes	Tetragonal
46	0.20	-	-	-	-	0.80	Yes	Tetragonal
47	-	-	-	-	0.15	0.85	Yes	Cubic
48	-	-	-	-	0.30	0.70	Yes	Cubic
49	-	0.15	-	-	-	0.85	Yes	Cubic
50	-	0.30	-	-	-	0.70	Yes	Cubic
51	-	-	0.05	-	-	0.95	Yes	Tetragonal
52	-	-	0.10	-	-	0.90	No	-
53	-	-	-	-	0.60	0.40	Yes	Cubic
54	0.10	-	-	-	0.60	0.30	Yes	Cubic
55	-	0.10	-	-	0.60	0.30	No	-
56	-	-	0.10	-	0.60	0.30	No	-
57	-	-	-	-	0.70	0.30	Yes	Cubic
58	0.10	-	-	-	0.70	0.20	No	-
59	-	0.10	-	-	0.70	0.20	No	-
60	-	-	0.10	-	0.70	0.20	No	-
61	0.05	-	-	-	0.05	0.90	Yes	Tetragonal
62	0.10	-	-	-	0.10	0.80	Yes	Tetragonal
63	0.15	-	-	-	0.15	0.70	Yes	Cubic
64	0.20	-	-	-	0.20	0.60	Yes	Cubic
65	0.25	-	-	-	0.25	0.50	Yes	Cubic

To collect the phase information, X-ray diffraction measurements (XRD) were performed in sintered samples using a Rigaku diffractometer, model Ultima IV, with Cu $K_{\alpha 1}$ radiation ($\lambda = 1.5406$ Å), operating at 40 kV and 40 mA, in step scan mode with a scanning step of 0.02° (4 s integrating time) over a 2θ range of 10° to 35°. The XRD data is given in Figure S1.

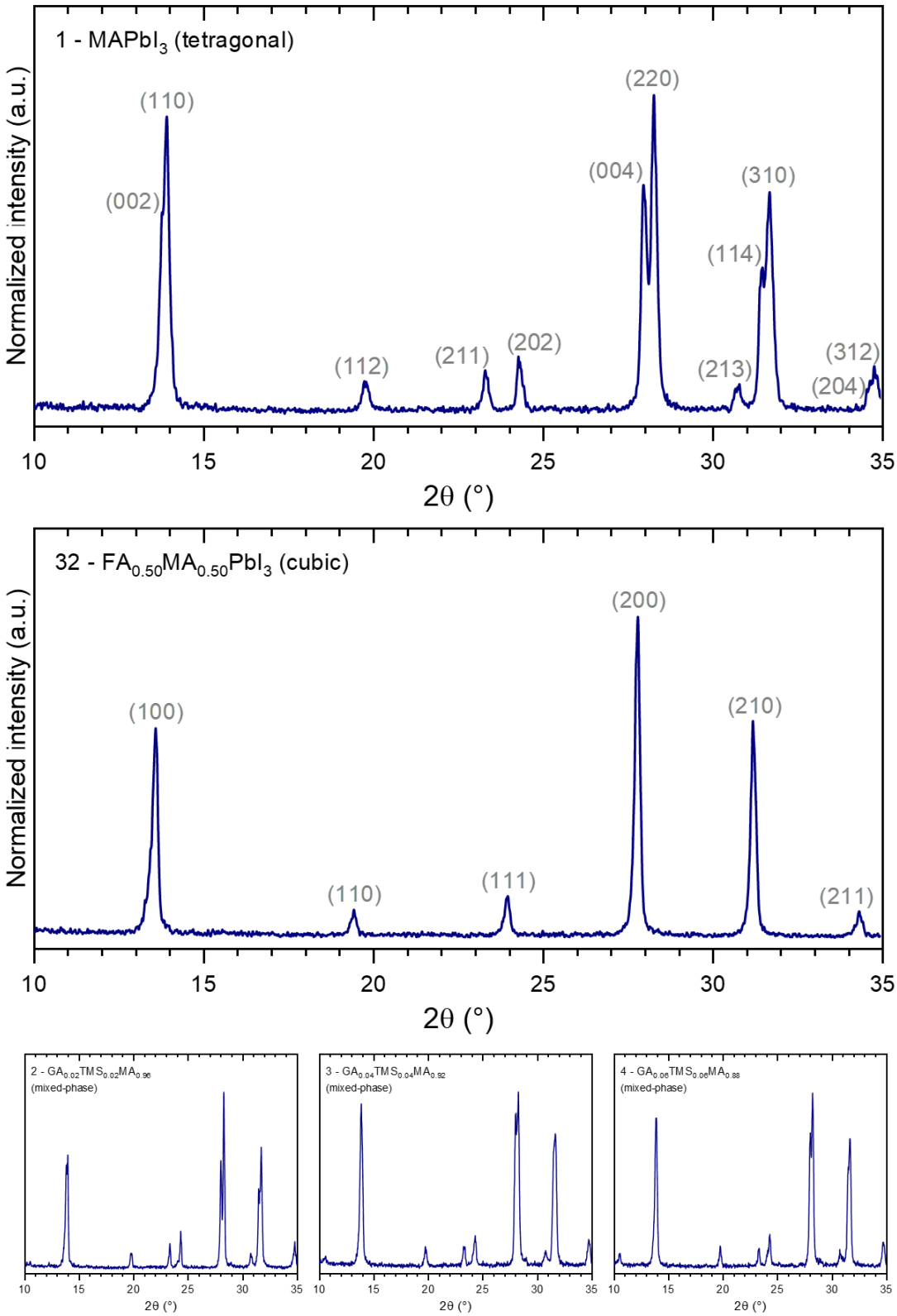


Figure S1 - XRD diffractograms of the synthesized perovskites. Index of tetragonal and cubic symmetries based on $I4cm$ and $Pm\bar{3}m$ space groups, respectively. Compositions with peaks other than the shown (compositions 1 and 32 used as references) were considered to have additional segregated phases, hence, are mixed-phase (that means, where perovskite and non-perovskite phases are present). Symmetry attribution of the single-phase perovskites was done through direct comparison. The main factor is the presence or absence of peaks at ~23.4 and ~30.6°, characteristic of the tetragonal phase.

Figure S1 - Continued.

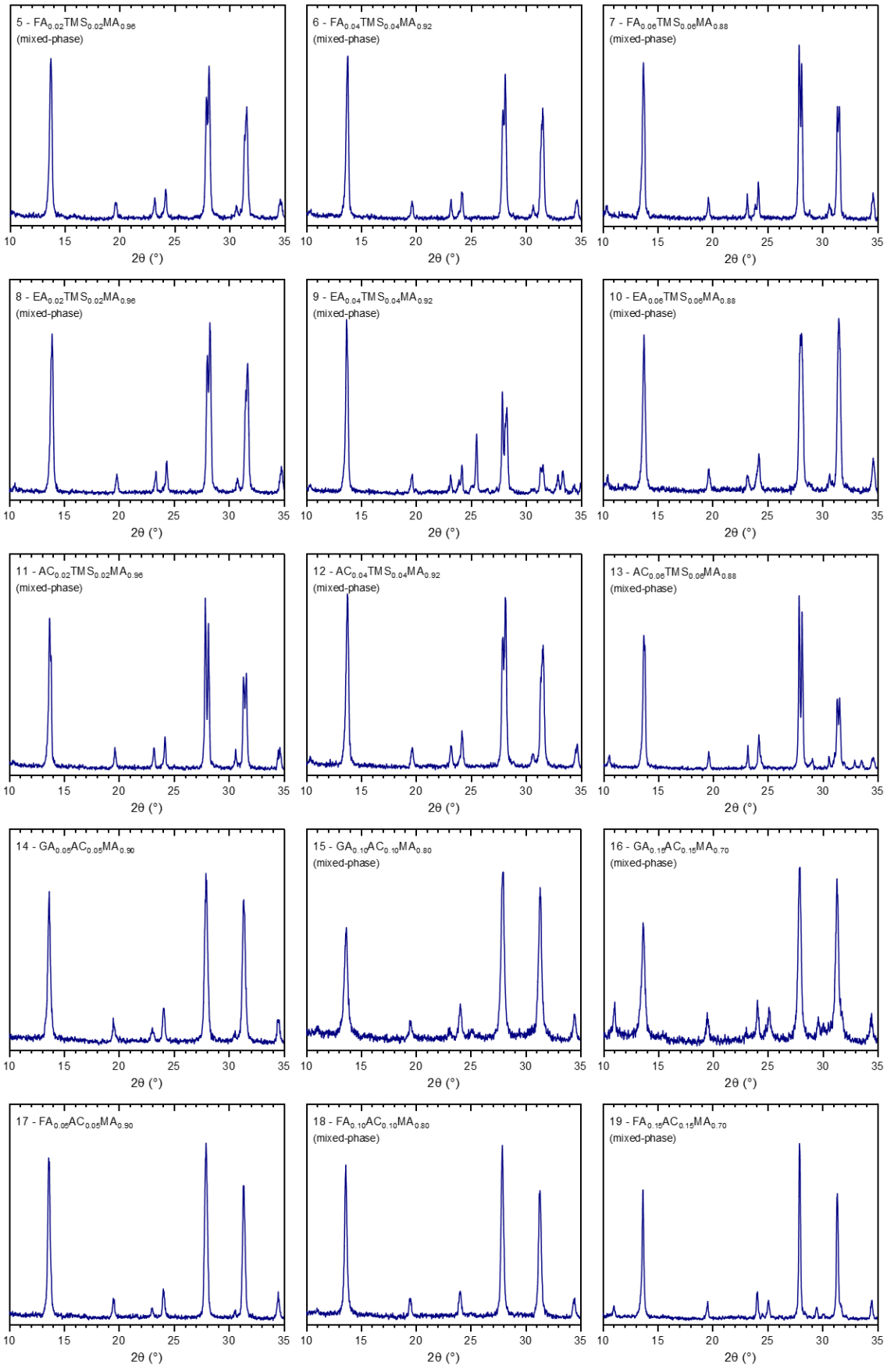


Figure S1 - Continued.

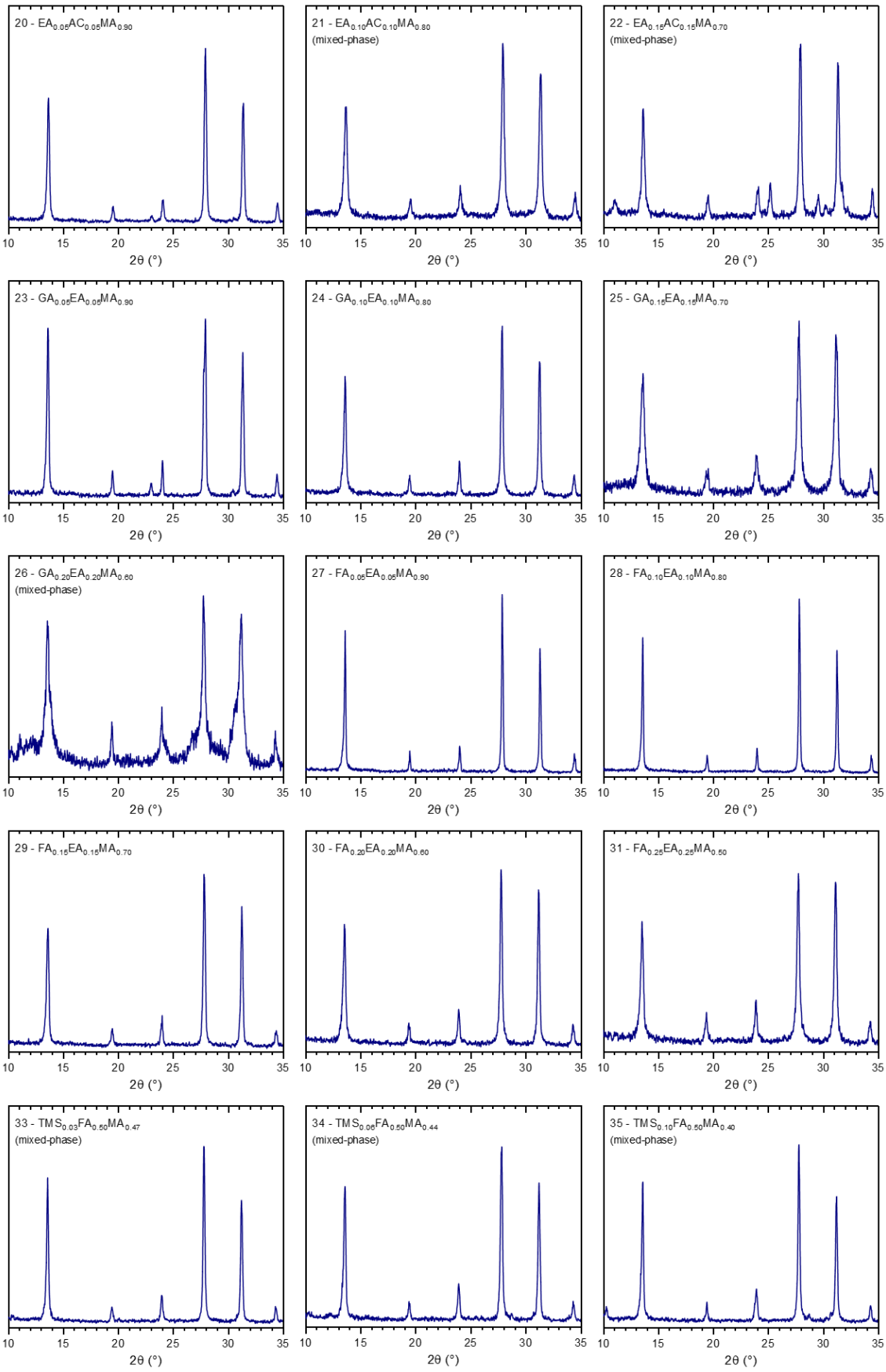


Figure S1 - Continued.

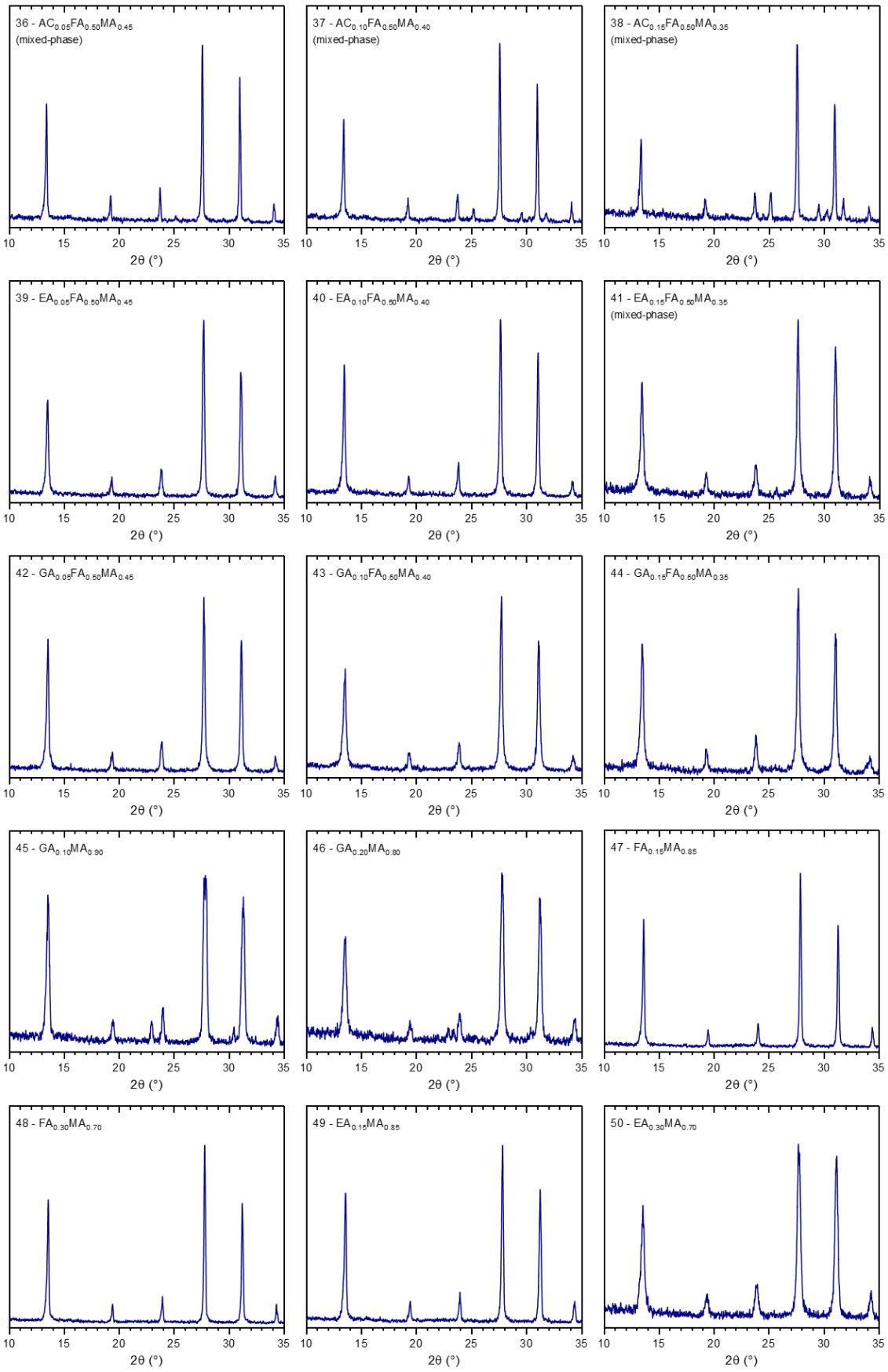
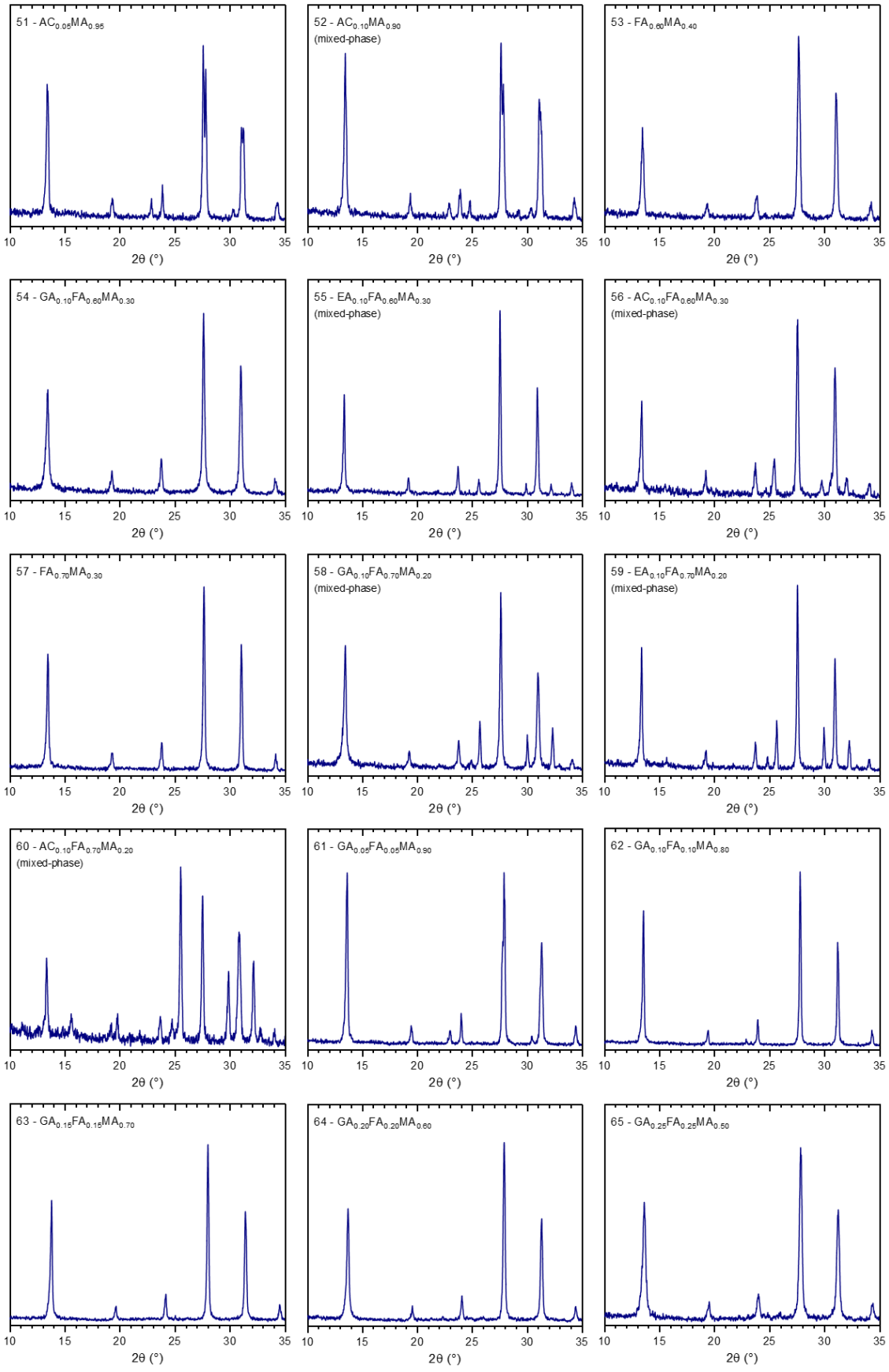


Figure S1 - Continued.



Supplementary Note 2: Experimental data and determination of properties

Differential scanning calorimetry (DSC) thermograms were acquired using 10 - 12 mg fragments of sintered pellets in hermetic aluminum crucibles with a DSC 25 apparatus (TA Instruments). Measurements were performed in a temperature range of -80 °C to 100 °C with a heating rate of 5 °C min^{-1} under a vacuum atmosphere. Results are given in Figure S2.

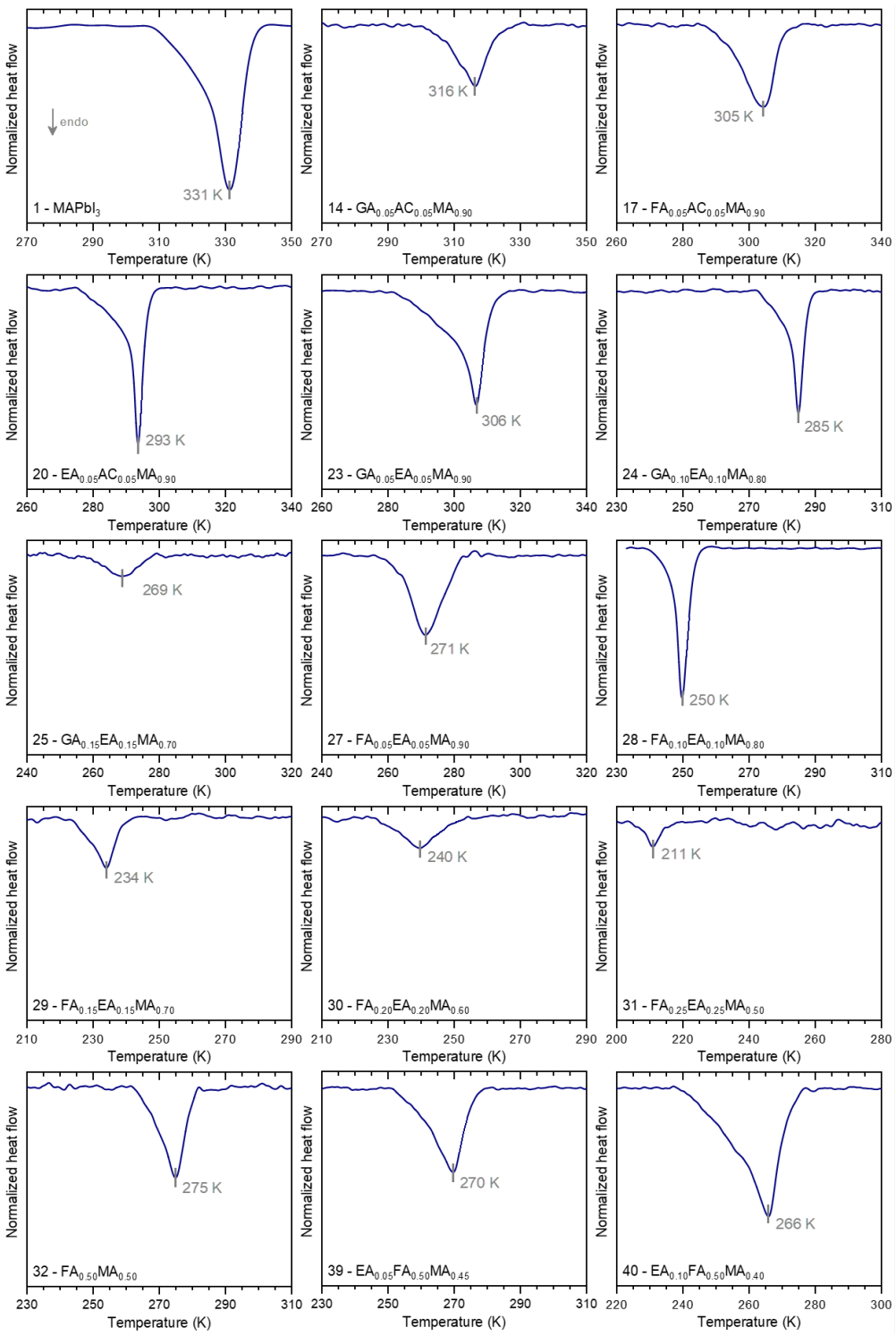
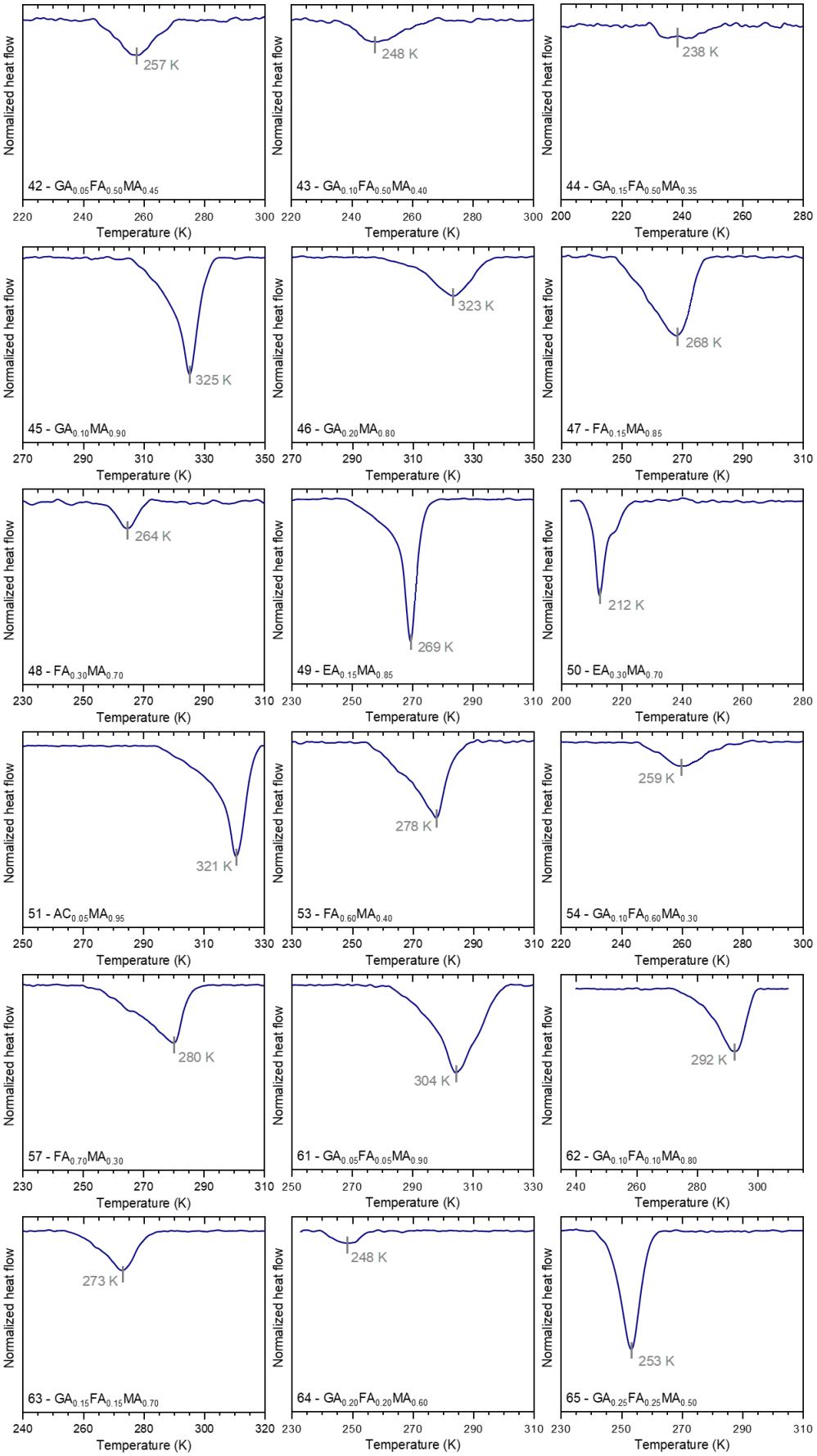


Figure S2 - DSC termograms of the single-phase compositions. For all curves, the transitions were considered at the peak of endotermic signals.

Figure S2 - Continued.



Diffuse UV-Vis reflectance spectroscopy measurements were collected using a Shimadzu UV-2600 spectrophotometer, with a range of 200 to 1400 nm, in low-speed mode with a data step of 1 nm. Based on these spectra, the band gap energies (E_g) were estimated through a linear fit of data in a Tauc plot (Figure S3) using the Kubelka-Munk-modified Tauc equation, given by $(F(R).h\nu)^{1/\gamma} = B(h\nu - E_g)$, where $h\nu$ is the incident photon energy, γ a factor that depends on the nature of the electronic transition, and $F(R)$ is the Kubelka-Munk function, given by $F(R) = (1-R)^2/2R$, where R is the reflectance of an infinitely thick sample [1,2]. We considered electronic transitions related to a direct band gap ($\gamma = 1/2$), as reported for halide perovskites [3].

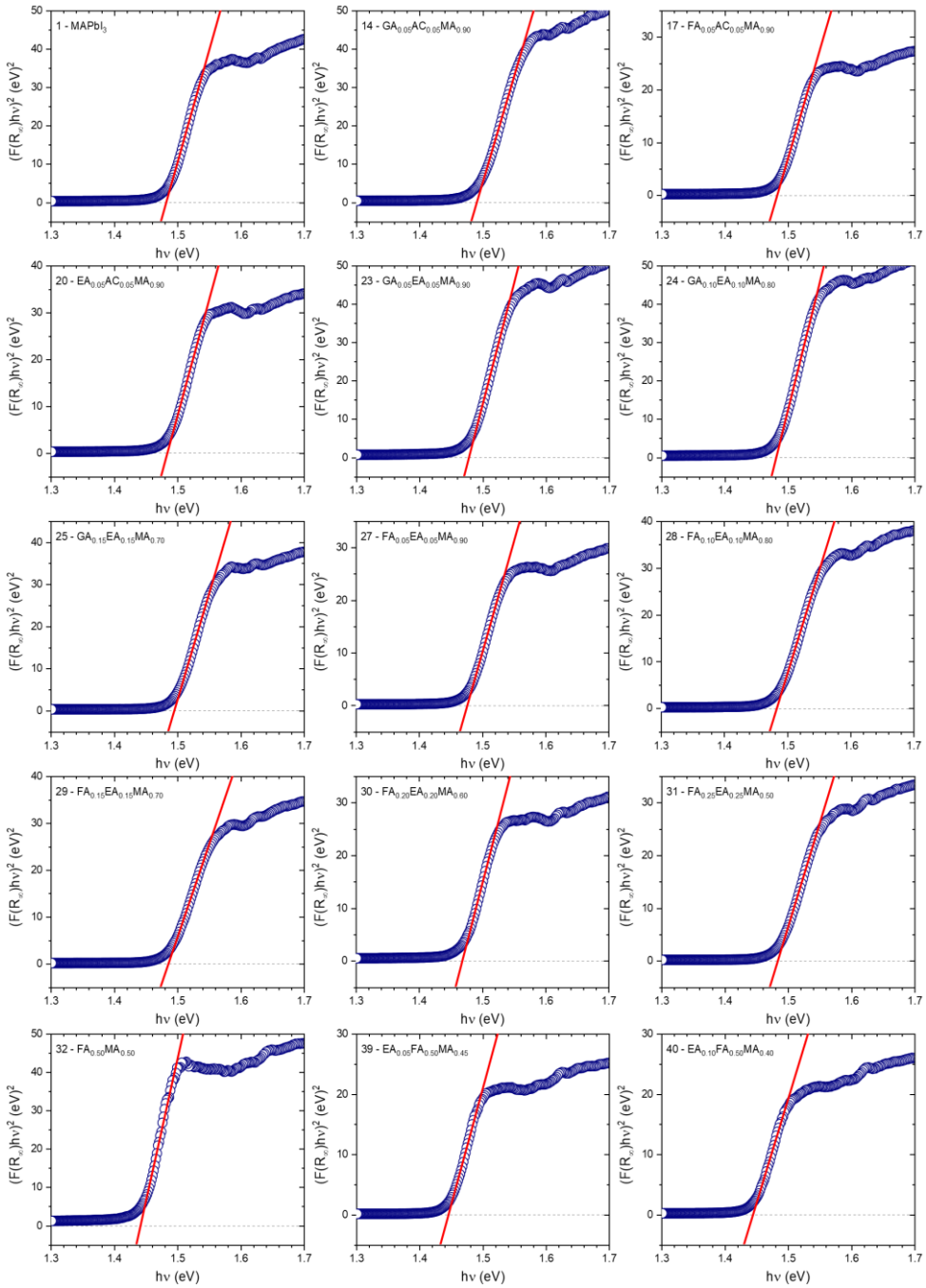
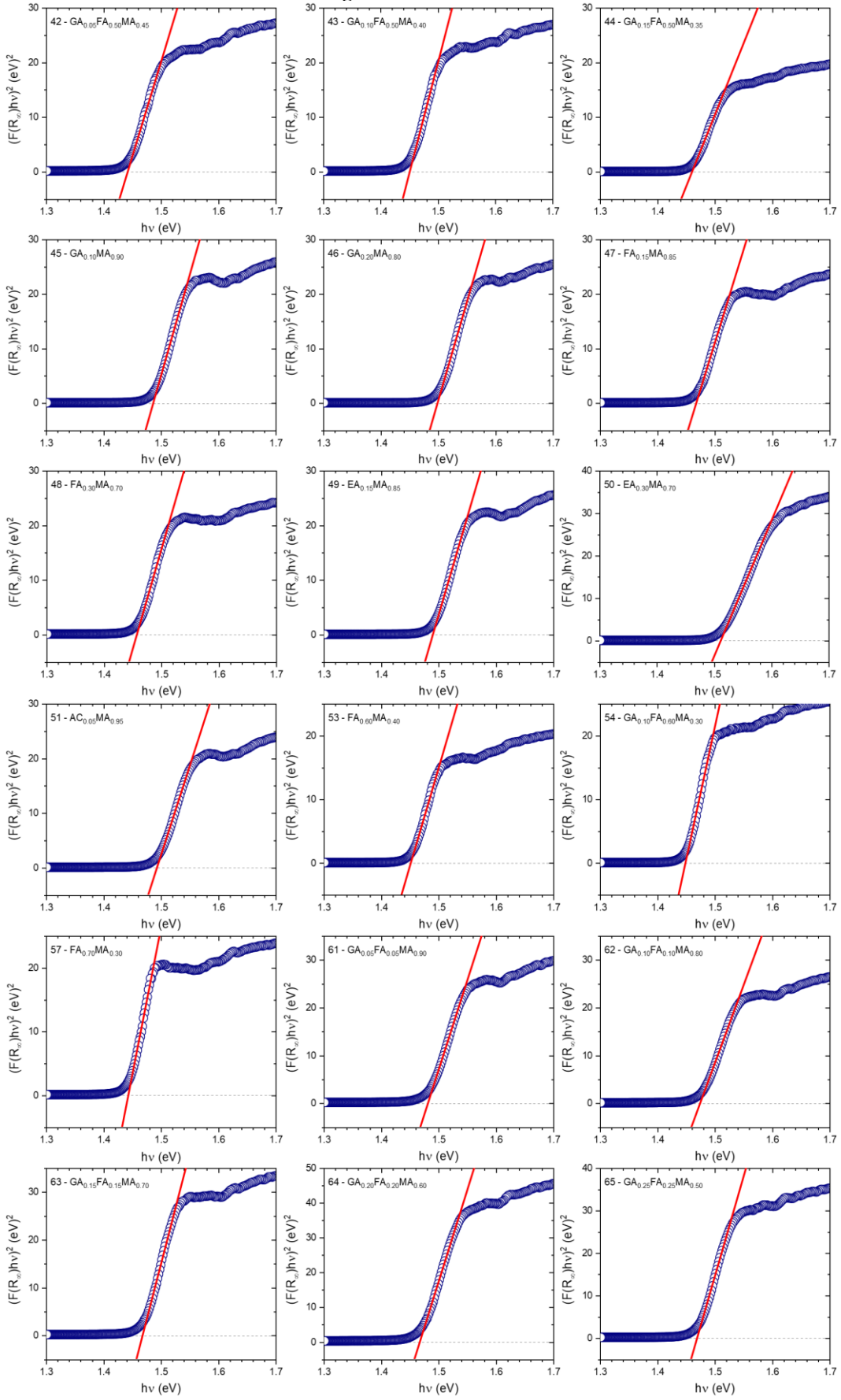


Figure S3 - Tauc plots and respective linear fits (red line) of the single-phase compositions.

Figure S3 - Continued.



Using the absorption data, we extracted the Urbach energies (E_U) using a fiting expression in the form $\alpha(E) = \alpha_0 \exp\left[(E - E_0)/E_U\right] + D$, where α_0 is a fitting parameter, $E = h\nu$ the photon energy, E_0 a material constant, and D an optional offset accounting for light scattering in the data [4]. In the present case, E_0 was considered to be band gap energies extracted previously. The experimental curves and fitting lines are given in Figure S4.

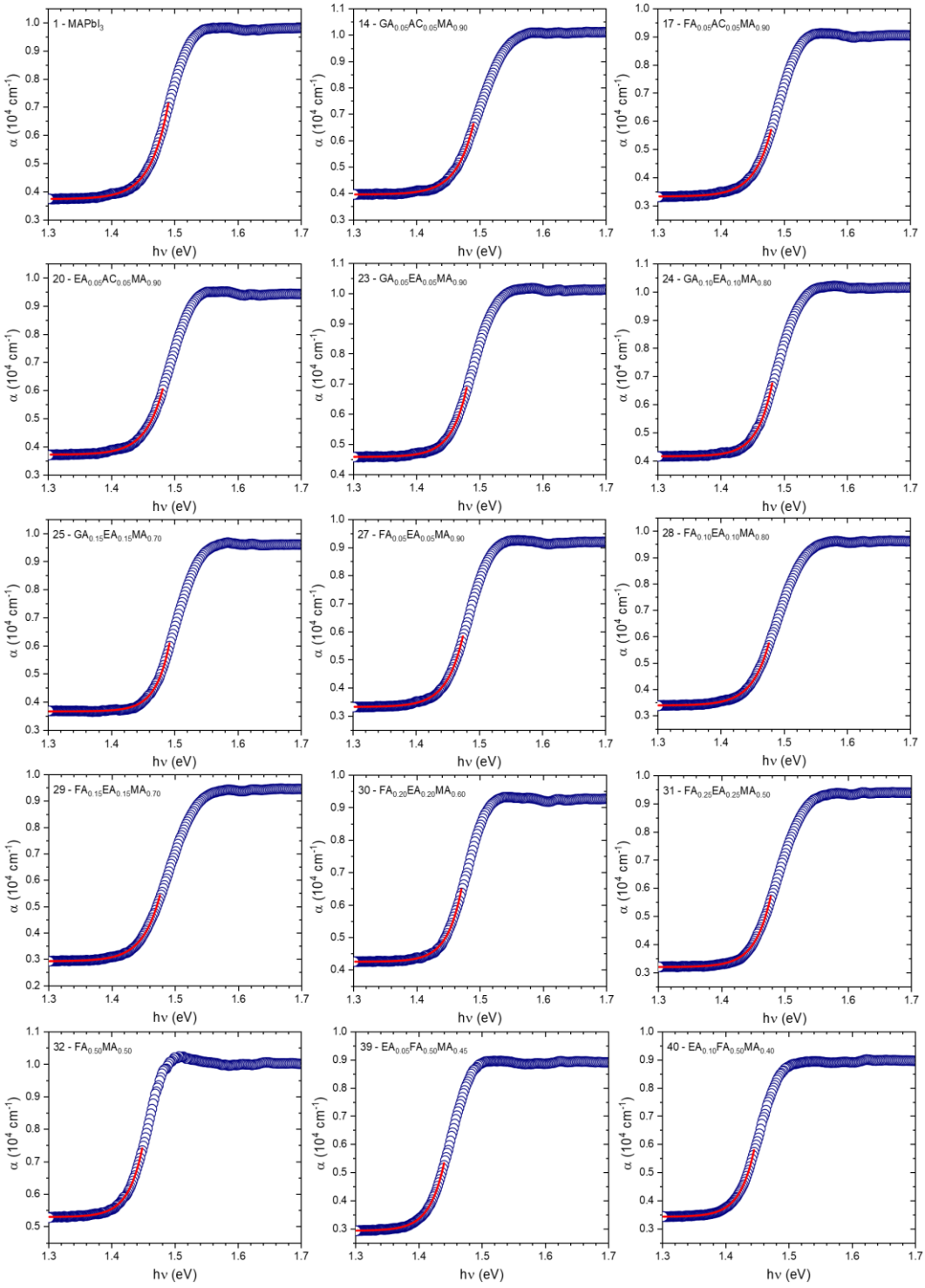
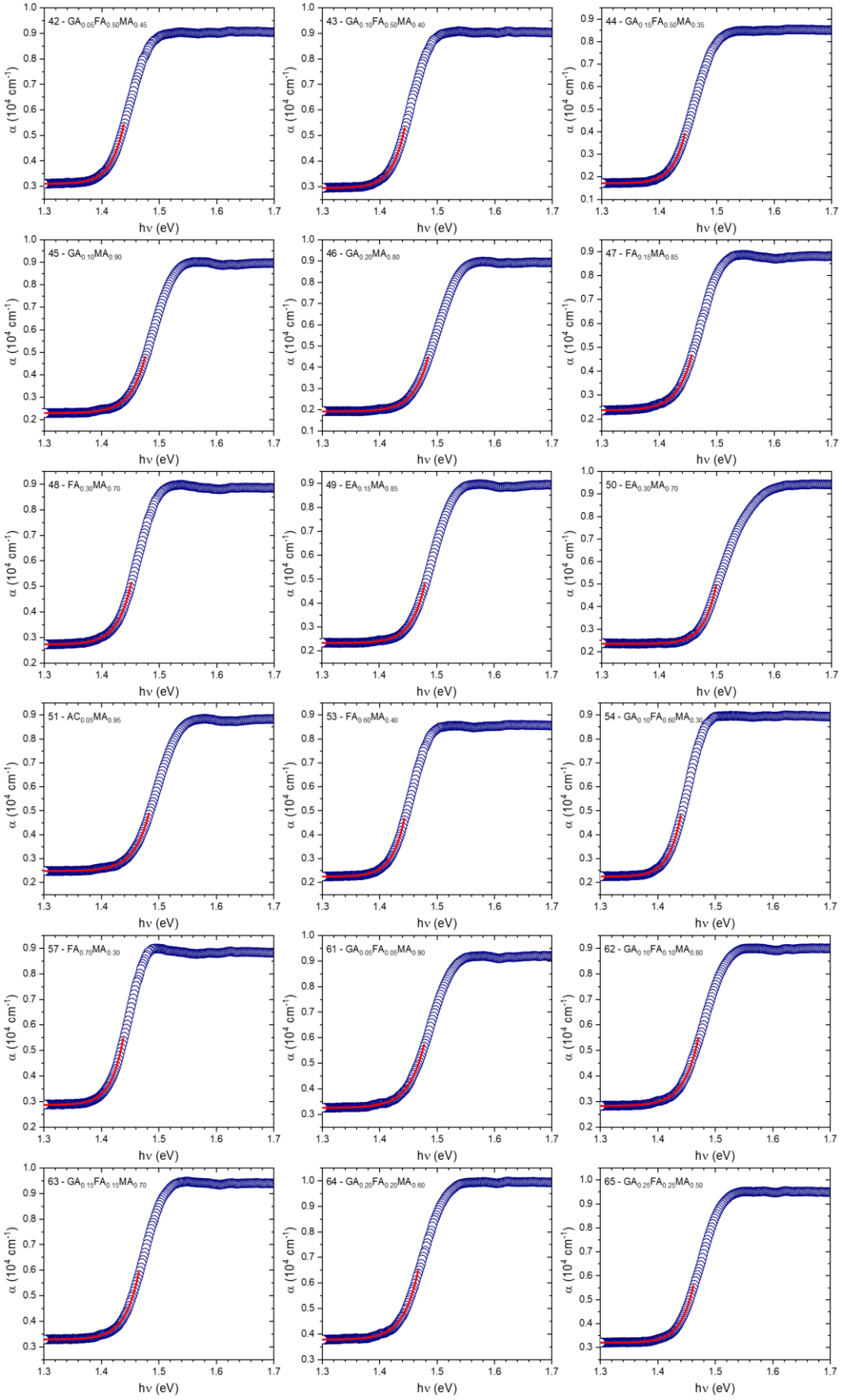


Figure S4 - Absorpion coefficient curves as a function of the photon energy for single-phase compositions. Red lines are the fits with $\alpha(E)$ equation.

Figure S4 - Continued.



For electrical measurements, nine circular gold electrodes of approximately 1 mm in diameter were sputtered onto the top of the perovskite pellets, while a gold electrode was deposited over the entire bottom surface of the sample. Steady-state current-voltage (*I-V*) data were collected in three randomly chosen electrodes using a Keithley 6517B electrometer in the dark at room temperature. Up-down cycle measurements were performed from ±6 V at a scan rate of 0.1 V s^{-1}, while forward measurements were conducted from 0 to +30 V at a scan rate of 0.5 V s^{-1}. Data collect in forward and cycle modes are given in Figures S5 and S6, respectively.

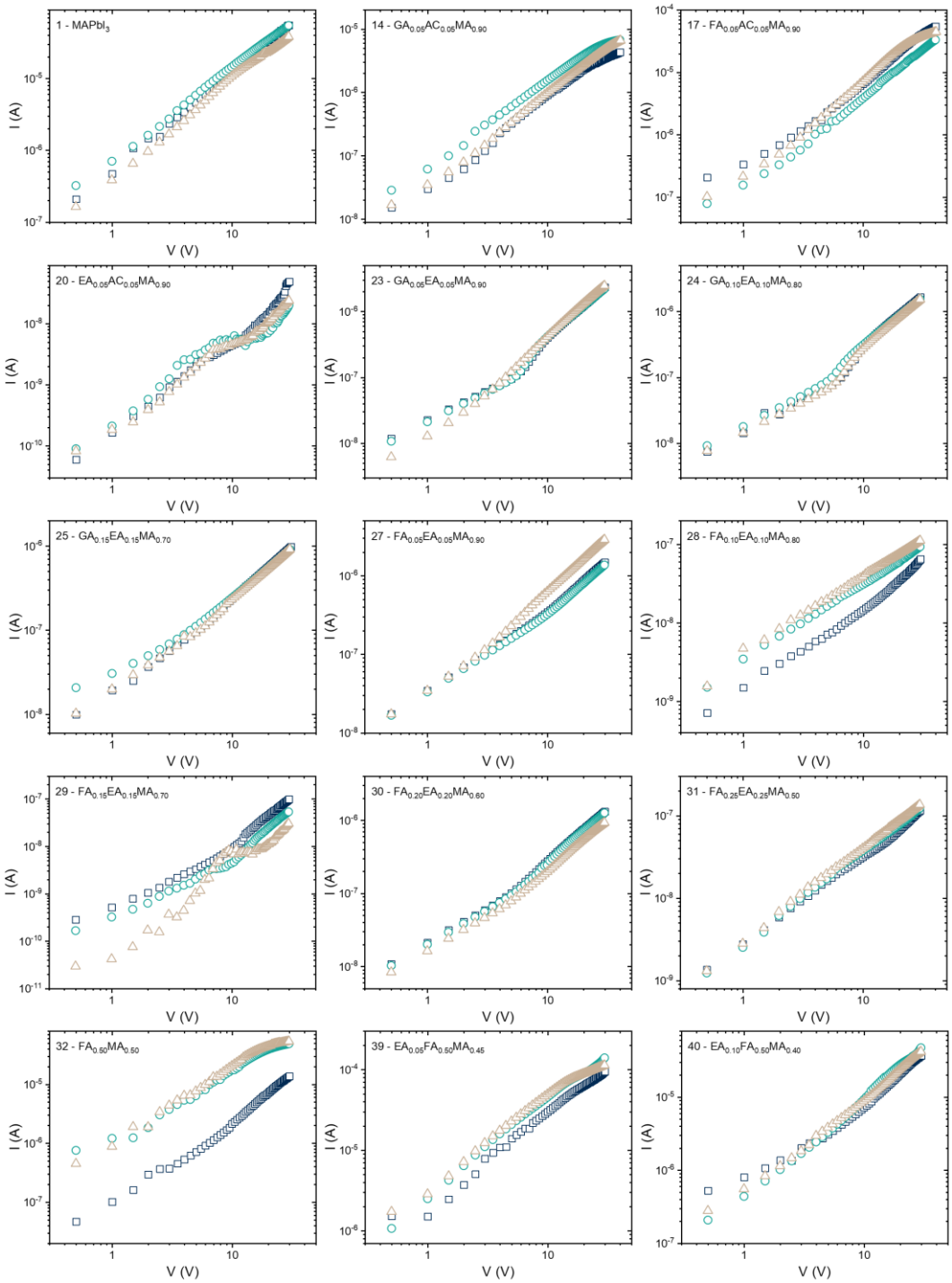
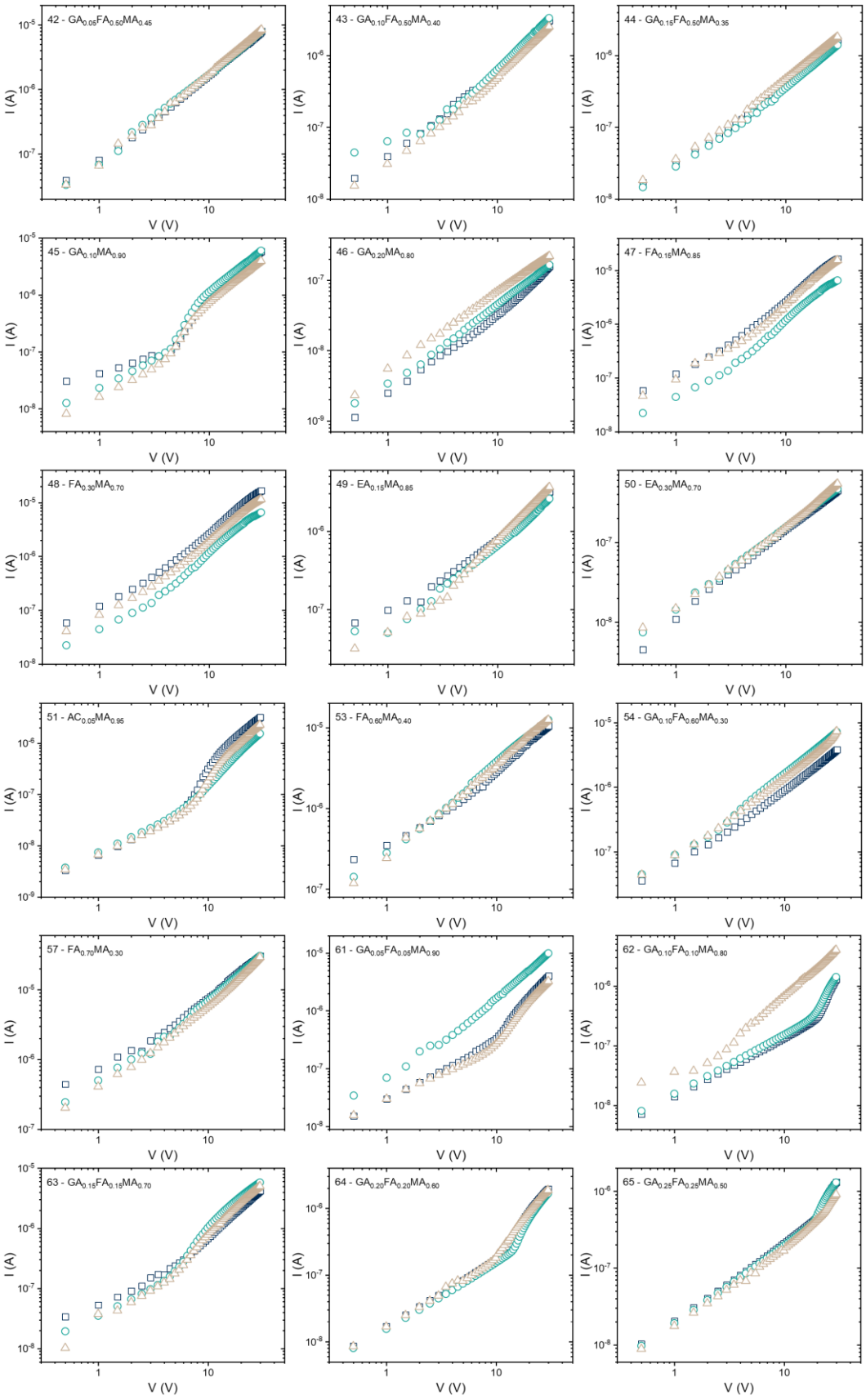


Figure S5 - Forward current-voltage data of single-phase compositions. The three colors are of each electrode.

Figure S5 - Continued.



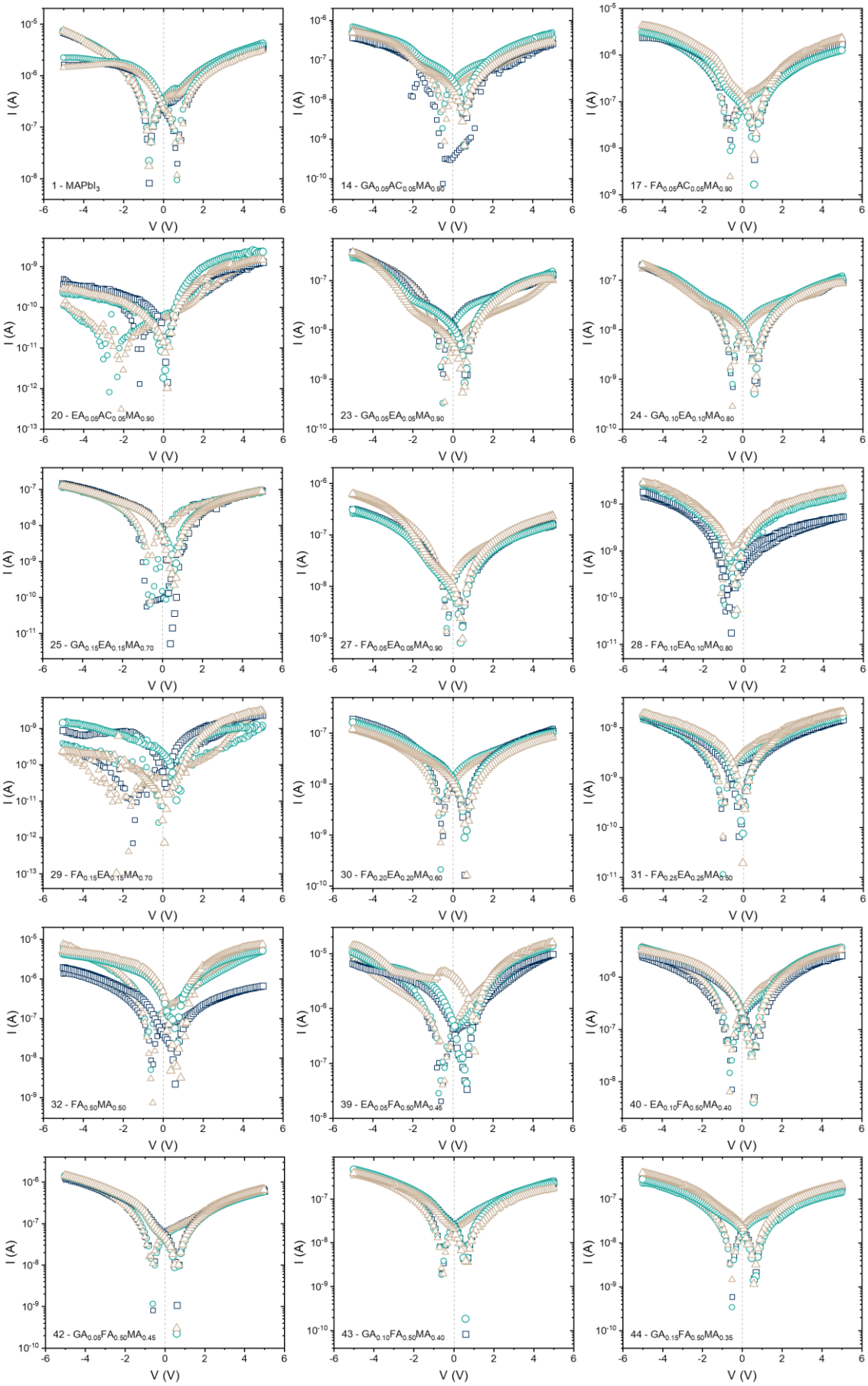
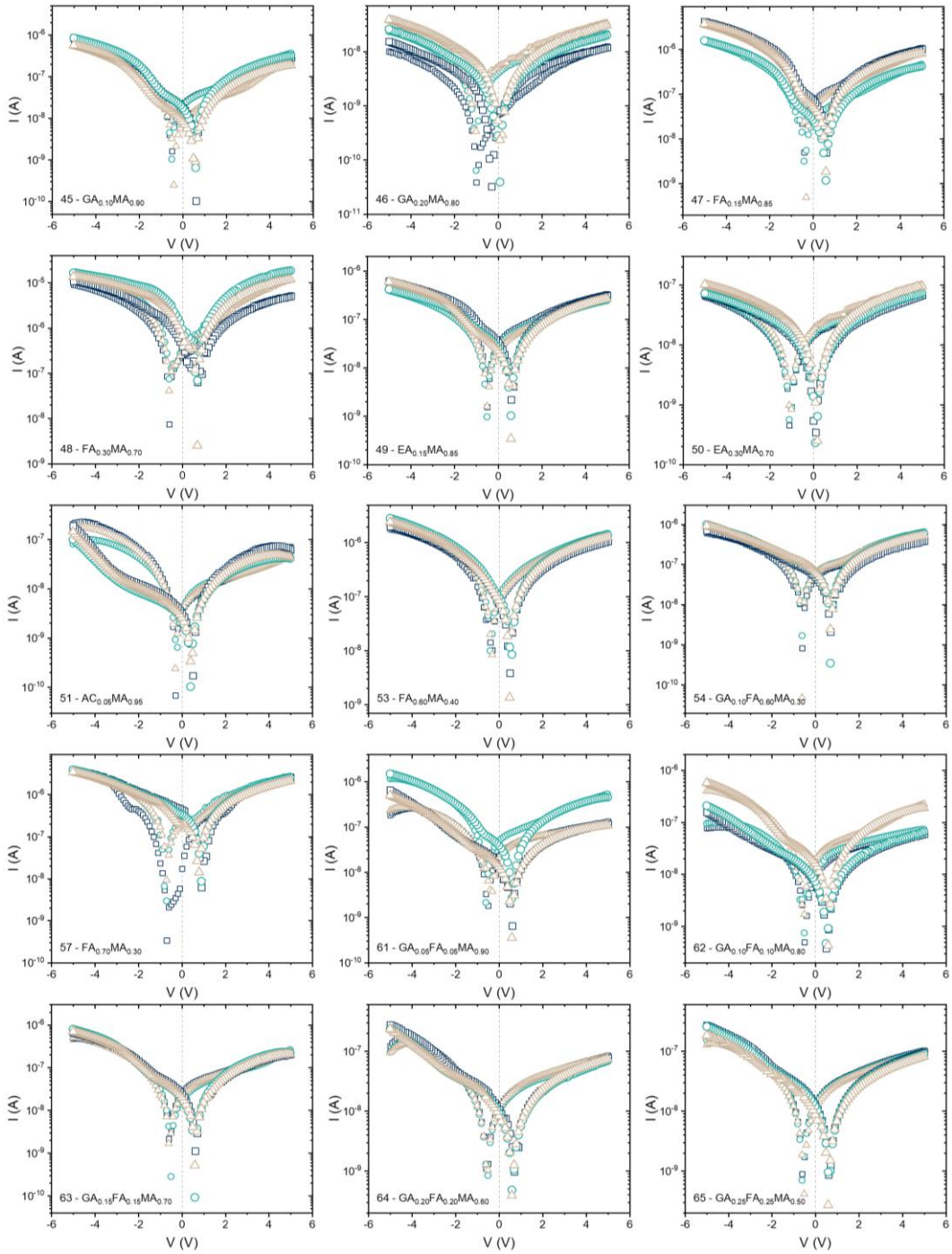


Figure S6 - Up-and-down current-voltage data of single-phase compositions. The three colors are of each electrode.

Figure S6 - Continued.



From the forward curves, electrical conductivities (σ_e) were obtained from linear fits in the low voltage regimes (below ~3 V), where $slope = I/V = 1/R$, where R is the measured resistance. Then, $\sigma_e = L/AR$, where L is the sample's thickness and A the electrode area. From the cycles, hysteresis values were estimated by the areas between up and down curves, similarly to [5]. The areas are normalized by the volumes $V = AL$ of charge transport. To obtain the A and L, micrographs of the samples were collected by scanning electron microscopy using a Zeiss scanning electron microscope model EVO LS15, operating with a tungsten beam under a voltage of 20 kV and by secondary electron analysis (SE mode). The micrographies are shown in Figures S7 and S8.

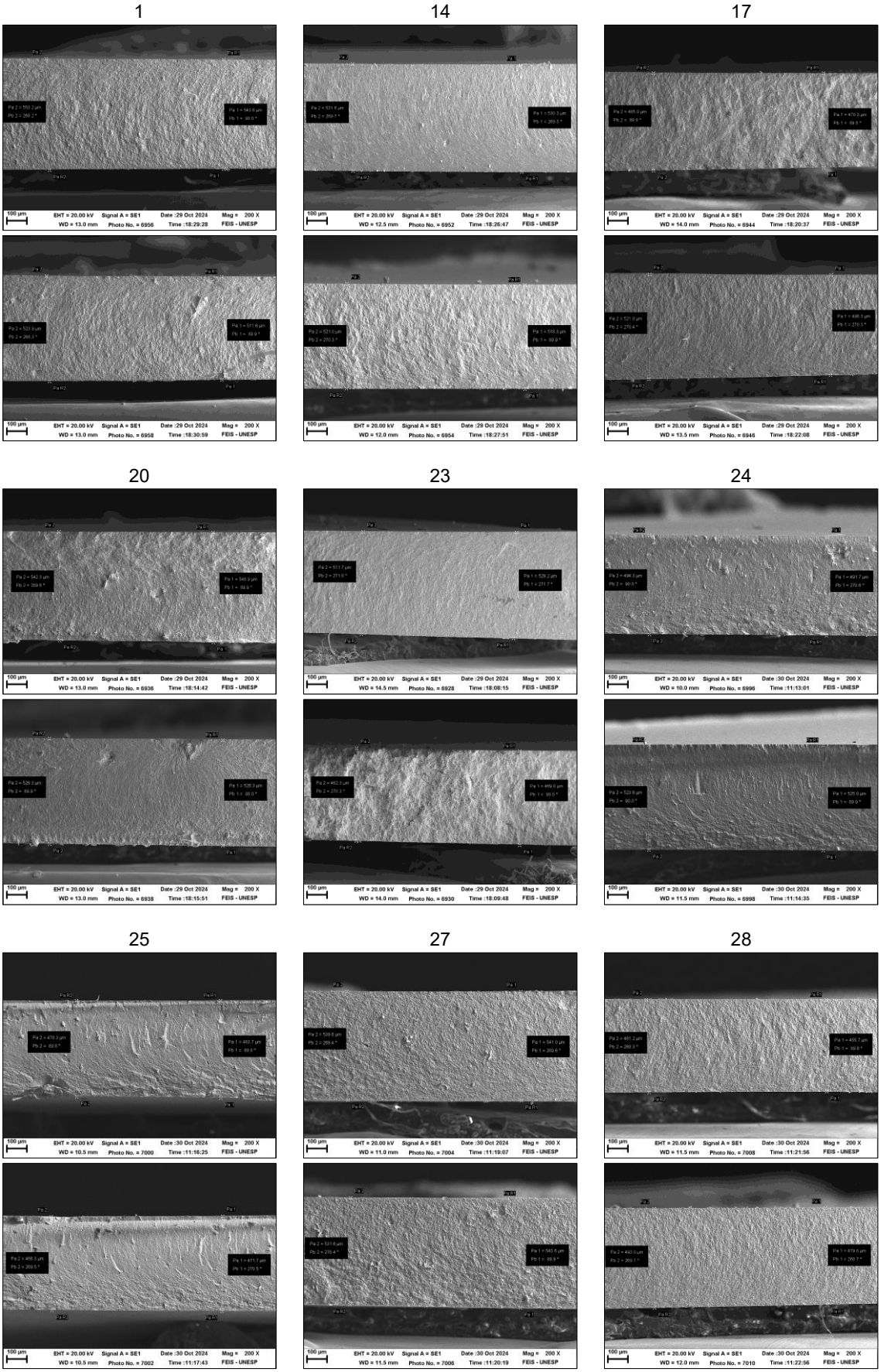
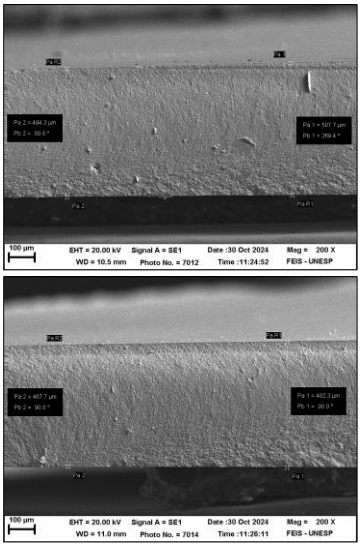


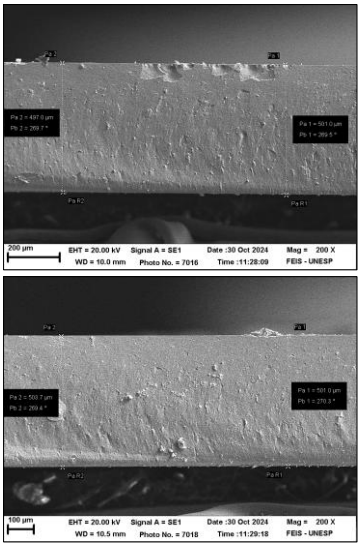
Figure S7 - Cross section scanning electron microscopy images of samples used to estimate the thicknesses. The final values for each sample are the mean of the four measurements shown in the images. The numbers above the images are the number of the compositions as shown in the Table S1.

Figure S7 - Continued.

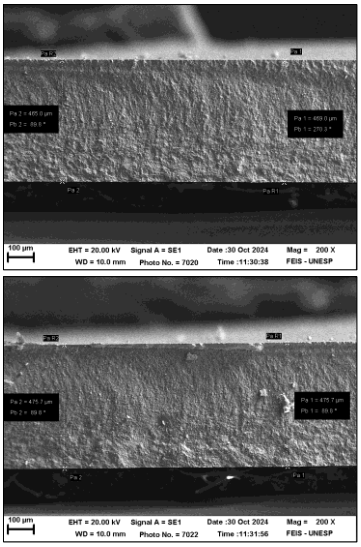
29



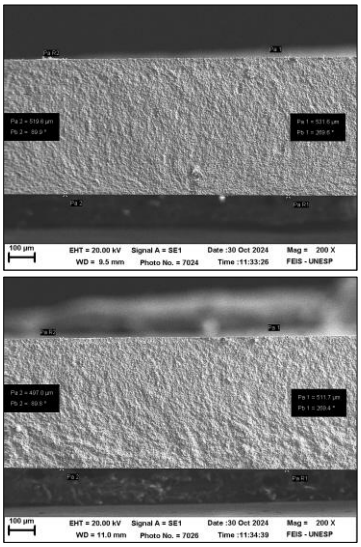
30



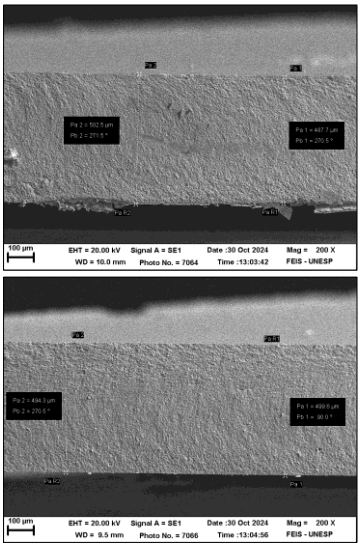
31



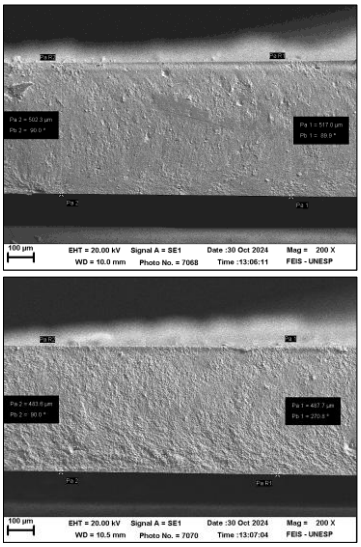
32



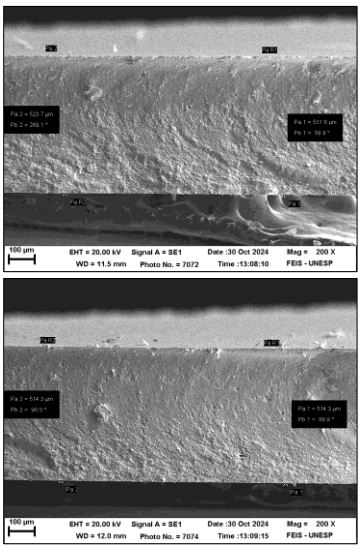
39



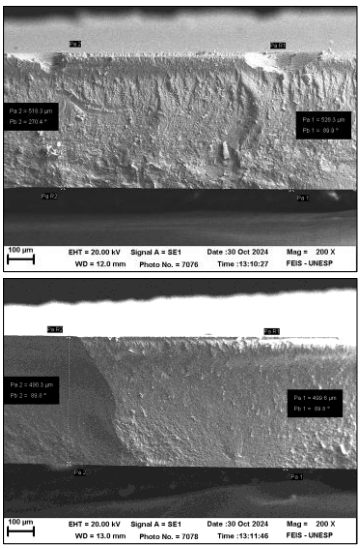
40



42



43



44

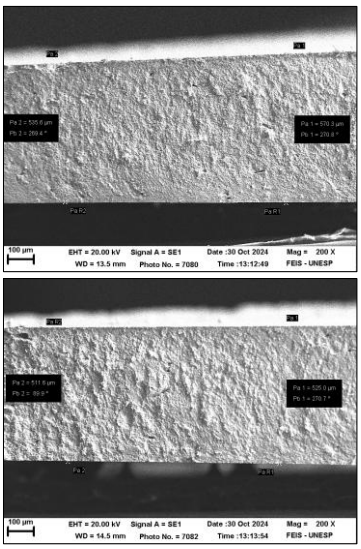
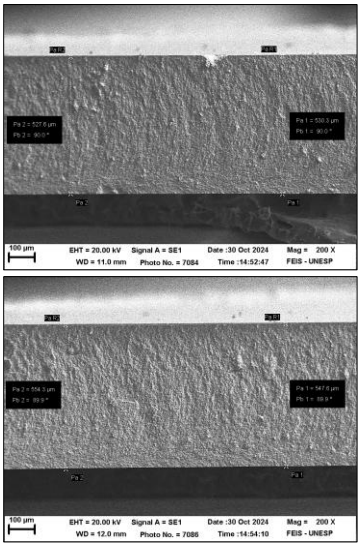
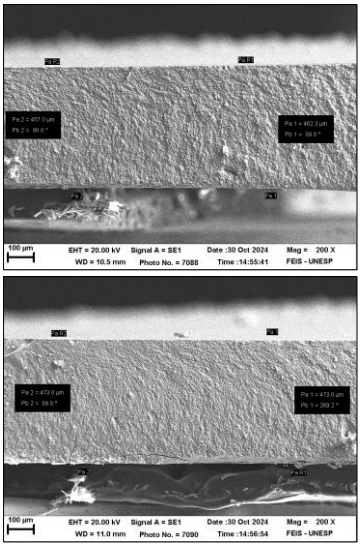


Figure S7 - Continued.

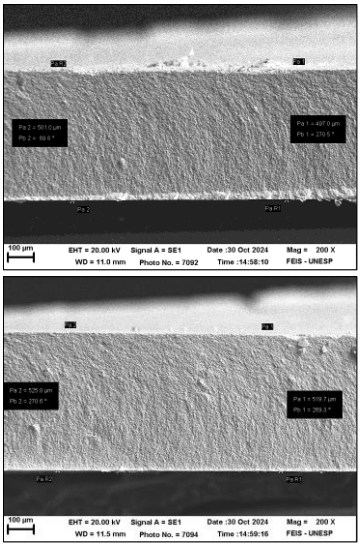
45



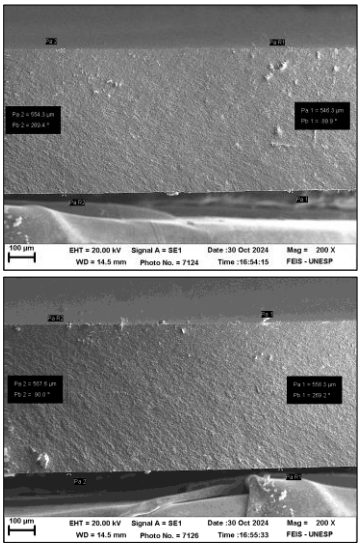
46



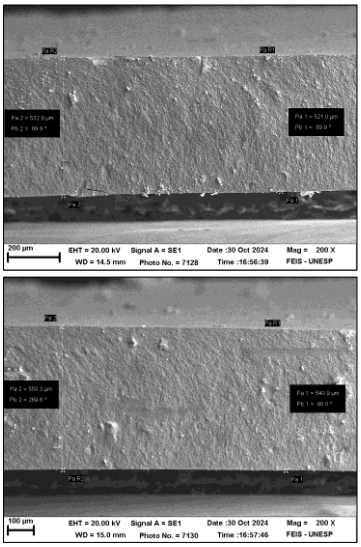
47



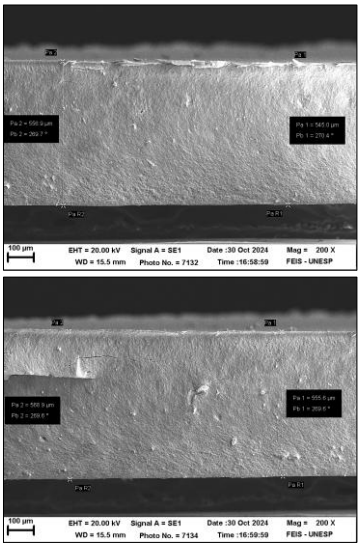
48



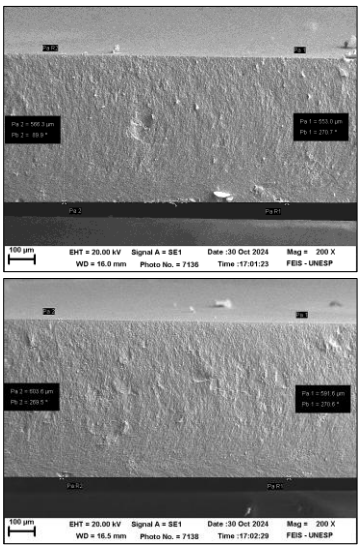
49



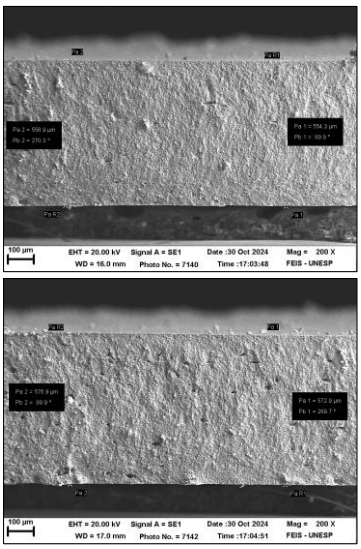
50



51



53



54

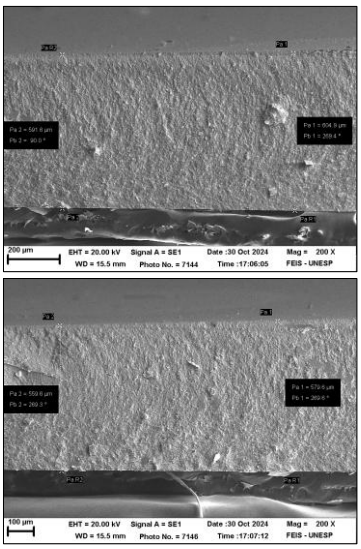
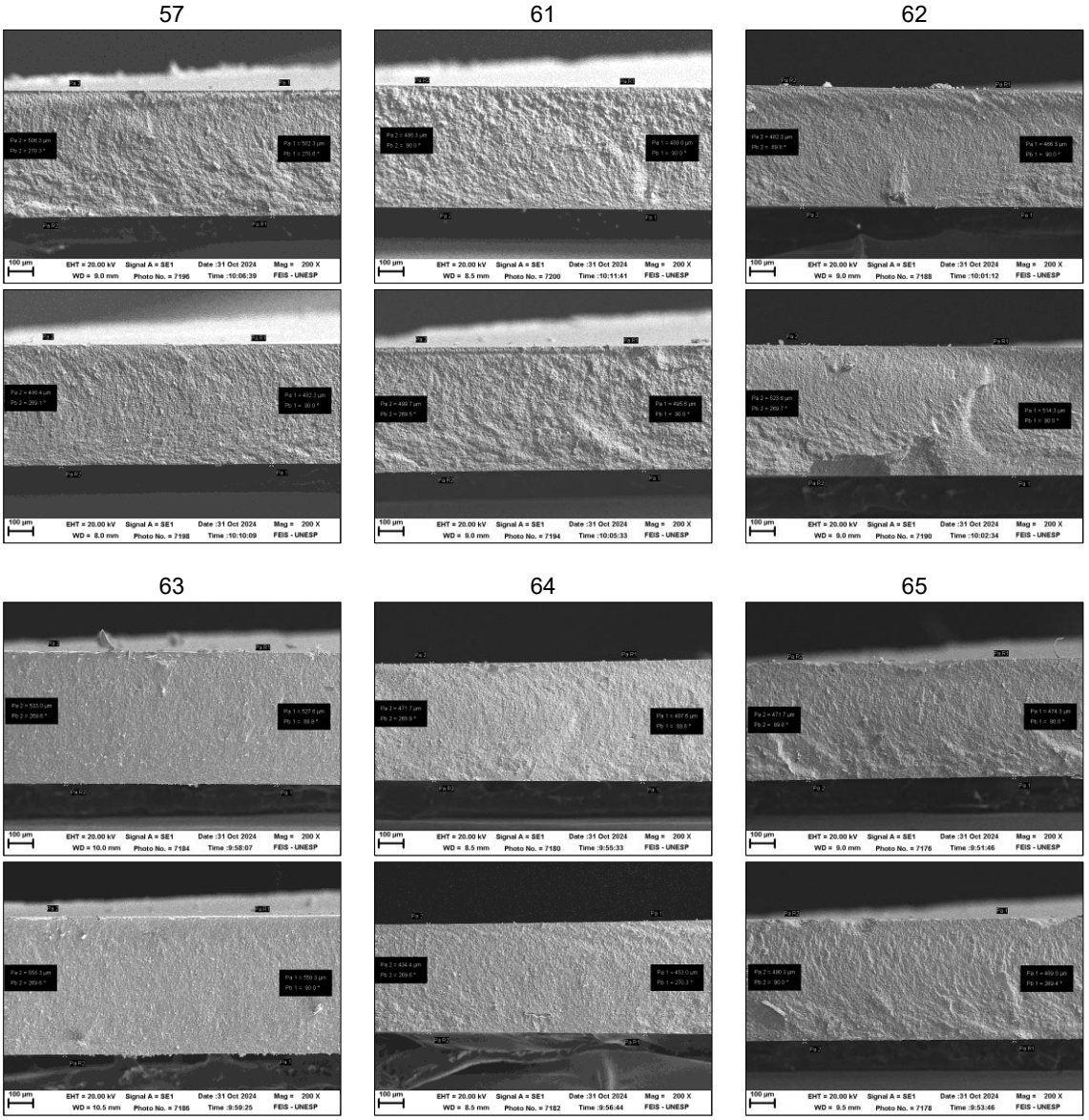


Figure S7 - Continued.



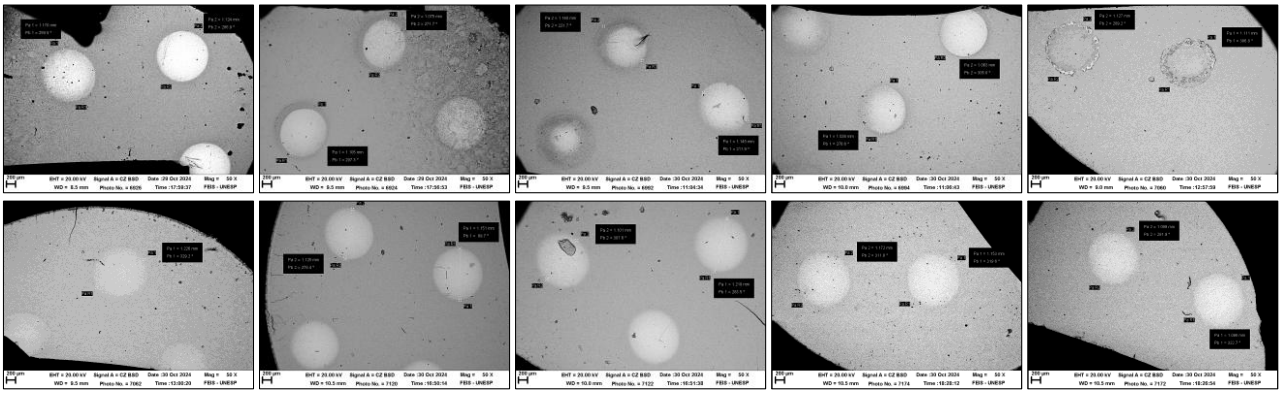


Figure S8 - Surface scanning electron microscopy images of random samples showing some electrodes and their approximate diameters. Since images were collected after the electrical measurements, some electrodes appear destroyed. Final electrode area for all samples were calculated from the mean values of the 19 measured diameters.

Supplementary Note 3: Effective descriptors and experimental properties

Table S2 - Summary of the single-phase compositions transformed into their effective descriptors and the respective measured properties. Deviations for band gap and Urbach energies are the fit residuals, while for conductivity and hysteresis are the standard deviations from three measurements. Values in bold were excluded in data selection step (see next section).

Composition number	Effective descriptors				T → C	Band gap (eV)		Urbach (meV)		Conductivity (S cm^{-1})		Hysteresis (W cm^{-3})	
	r_{eff} (pm)	μ_{eff} (D)	NH_{eff}	ΔS_c (R)	temp. (K)	value	dev.	value	dev.	value	dev.	value	dev.
1	217.0	2.69	3.00	0.000	331	1.482	0.012	**28.46**	**0.21**	3.36E-06	5.45E-07	**4.04E-02**	**3.75E-03**
14	223.1	2.52	3.20	0.079	316	**1.491**	**0.011**	**27.43**	**0.26**	2.38E-07	5.67E-08	9.81E-04	8.55E-05
17	221.8	2.53	3.10	0.079	305	1.483	0.012	27.11	0.17	1.21E-06	2.53E-07	**5.21E-03**	**5.02E-04**
20	222.9	2.77	3.05	0.079	293	1.484	0.013	**27.71**	**0.20**	**1.16E-09**	**1.50E-10**	**3.87E-06**	**3.42E-07**
23	222.9	2.67	3.15	0.079	306	1.478	0.008	24.48	0.23	9.15E-08	1.42E-08	8.30E-04	8.98E-05
24	228.8	2.65	3.30	0.128	285	**1.481**	**0.007**	23.70	0.22	7.72E-08	8.59E-09	3.17E-04	2.94E-05
25	234.7	2.63	3.45	0.164	269	1.495	0.012	**22.16**	**0.25**	9.69E-08	1.00E-08	9.81E-05	8.63E-06
27	221.7	2.68	3.05	0.079	271	1.477	0.010	25.94	0.17	**1.82E-07**	**1.60E-08**	3.92E-04	3.37E-05
28	226.3	2.67	3.10	0.128	250	1.484	0.009	27.04	0.21	**1.39E-08**	**3.46E-09**	**3.17E-05**	**3.04E-06**
29	231.0	2.66	3.15	0.164	234	**1.486**	**0.009**	**27.44**	**0.26**	**1.61E-09**	**5.92E-10**	**7.04E-06**	**6.62E-07**
30	235.6	2.65	3.20	0.190	**240**	**1.469**	**0.010**	22.94	0.25	9.12E-08	9.85E-09	1.15E-04	9.90E-06
31	240.3	2.64	3.25	0.208	211	1.484	0.009	**24.88**	**0.30**	1.60E-08	1.61E-09	**1.46E-05**	**1.27E-06**
32	235.0	1.46	3.50	0.139	275	1.442	0.021	22.79	0.16	4.87E-06	1.88E-06	2.03E-02	1.87E-03
39	237.9	1.57	3.50	0.171	**270**	1.446	0.015	23.13	0.14	1.33E-05	2.30E-06	**2.14E-02**	**1.90E-03**
40	240.7	1.68	3.50	0.189	**266**	**1.445**	**0.013**	22.86	0.13	2.92E-06	3.07E-07	7.48E-03	6.81E-04
42	238.1	1.32	3.65	0.171	257	1.442	0.019	21.93	0.18	5.14E-07	4.93E-08	6.85E-04	5.98E-05
43	241.1	1.19	3.80	0.189	248	1.451	0.011	21.85	0.17	**2.13E-07**	**2.63E-08**	2.33E-04	2.16E-05
44	244.2	1.05	3.95	0.200	**238**	**1.460**	**0.016**	21.36	0.18	1.68E-07	1.89E-08	1.79E-04	1.75E-05
45	223.1	2.42	3.30	0.065	325	1.487	0.011	25.27	0.19	1.27E-07	2.36E-08	9.12E-04	8.12E-05
46	229.2	2.15	3.60	0.100	**323**	**1.499**	**0.015**	25.26	0.21	**1.85E-08**	**4.09E-09**	3.51E-05	3.18E-06
47	222.4	2.32	3.15	0.085	**268**	1.468	0.012	25.29	0.14	4.89E-07	1.23E-07	1.29E-03	1.18E-04
48	227.8	1.95	3.30	0.122	264	1.457	0.015	24.28	0.11	4.58E-07	1.14E-07	2.36E-02	2.06E-03
49	225.6	3.03	3.00	0.085	269	1.490	0.014	**23.49**	**0.20**	3.41E-07	4.31E-08	2.03E-04	1.88E-05
50	234.1	3.37	3.00	0.122	**212**	1.512	0.008	21.74	0.23	8.03E-08	7.40E-09	9.47E-05	8.28E-06
51	220.0	2.66	3.05	0.040	321	**1.493**	**0.011**	26.16	0.17	**3.95E-08**	**4.02E-09**	**5.87E-04**	**5.55E-05**
53	238.6	1.21	3.60	0.135	278	1.452	0.018	21.66	0.16	**1.58E-06**	**1.40E-07**	1.49E-03	1.31E-04
54	244.7	0.94	3.90	0.180	259	1.449	0.017	20.89	0.19	5.22E-07	7.09E-08	5.45E-04	5.05E-05
57	242.2	0.96	3.70	0.122	280	1.443	0.017	21.11	0.18	2.51E-06	3.35E-07	3.99E-03	3.58E-04
61	221.9	2.43	3.20	0.079	304	1.481	0.010	27.06	0.23	1.57E-07	2.45E-08	3.33E-03	2.88E-04
62	226.7	2.17	3.40	0.128	292	1.474	0.013	25.74	0.24	9.34E-08	2.14E-08	1.11E-03	1.13E-04
63	231.6	1.92	3.60	0.164	273	1.468	0.013	23.87	0.19	2.00E-07	3.15E-08	1.52E-03	1.38E-04
64	236.4	1.66	3.80	0.190	**248**	1.467	0.011	**25.57**	**0.22**	7.36E-08	7.58E-09	1.37E-03	1.36E-04
65	241.3	1.40	4.00	0.208	253	1.469	0.012	22.36	0.18	8.94E-08	8.95E-09	6.57E-04	5.78E-05

Supplementary Note 4: Statistical approach

Briefly, the effective descriptors of the compositions were used as the independent variables, while each one the the properties are the dependent variables. To perform the multivariate linear regression of data, the 33 points (compositions) were used as independent variables normalized using the z-score normalization. The resulting equation in the form $P_n = k_n + a_n r_{eff} + b_n \mu_{eff} + c_n \mathrm{NH}_{eff} + d_n \Delta S_c$ was used to calculate the property value for the compositions using the obtained k_n, a_n, b_n, c_n, and d_n coefficients. The experimental *versus* calculated values were plotted and a linear fit with slope equal to 1 was performed. To select the 25 data, the eight values farther from the $x = y$ line were removed. Then, the multivariate linear regression of data, calculation of properties, and linear fit procedure was redone. The results are given in Figure S9. For conductivity and hysteresis, we used the logarithm of values since they vary by several orders of magnitude.

It is important to mention that, given that a high degree of variability (low R^2) was present when using 33 points, we were concerned that at what point the descriptors were not randomly explaining the results. This could very much be a possibility in a rather small data bank, like the one we have. Hence, when excluding the eight points, we wanted to obtain better R^2 while also observing/confirming how strongly the descriptors affected the measured properties. Since the exclusion of the eight points did not change much the normalized coefficients (Figure S9) and lm.beta values (Figure S10), it means that the individual descriptors do not randomly affect the measured properties.

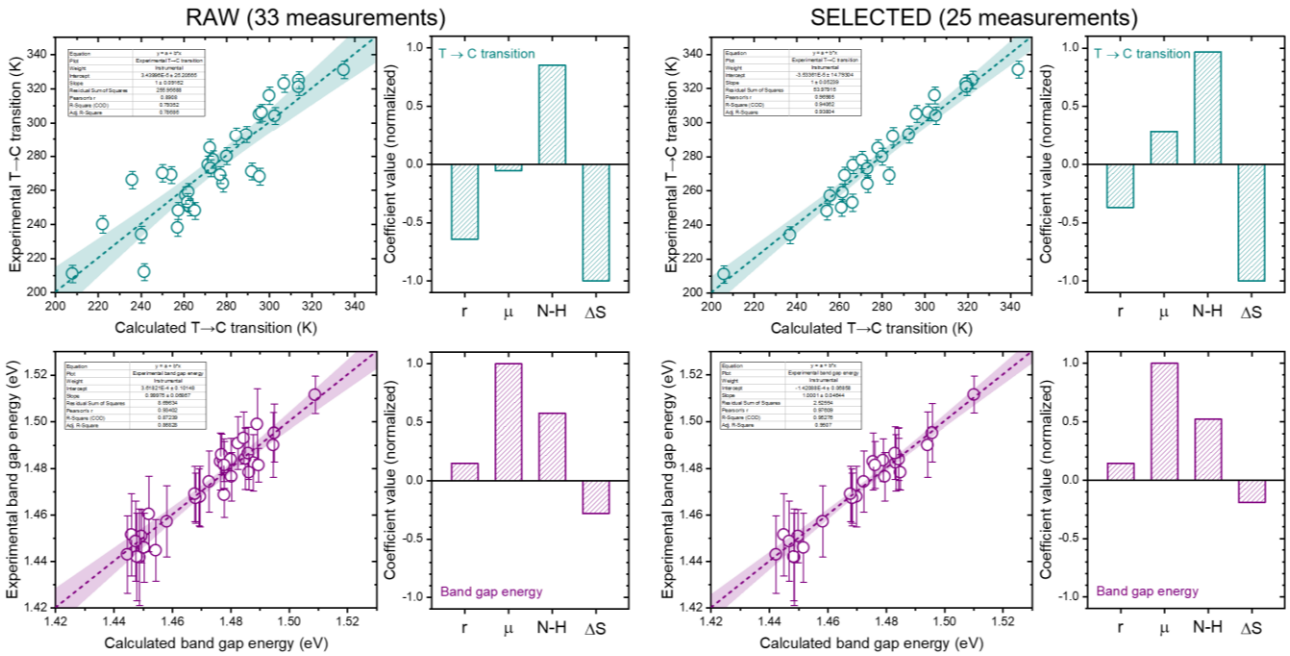


Figure S9 - Experimental versus calculated properties from 33 and 25 data. Shaded areas are the 95% confidence level bands. Noteworthy, relative coefficient values remain almost unchanged from raw to selected data, meaning that the removed data are not a source of systematic errors whatsoever.

Figure S9 - Continued.

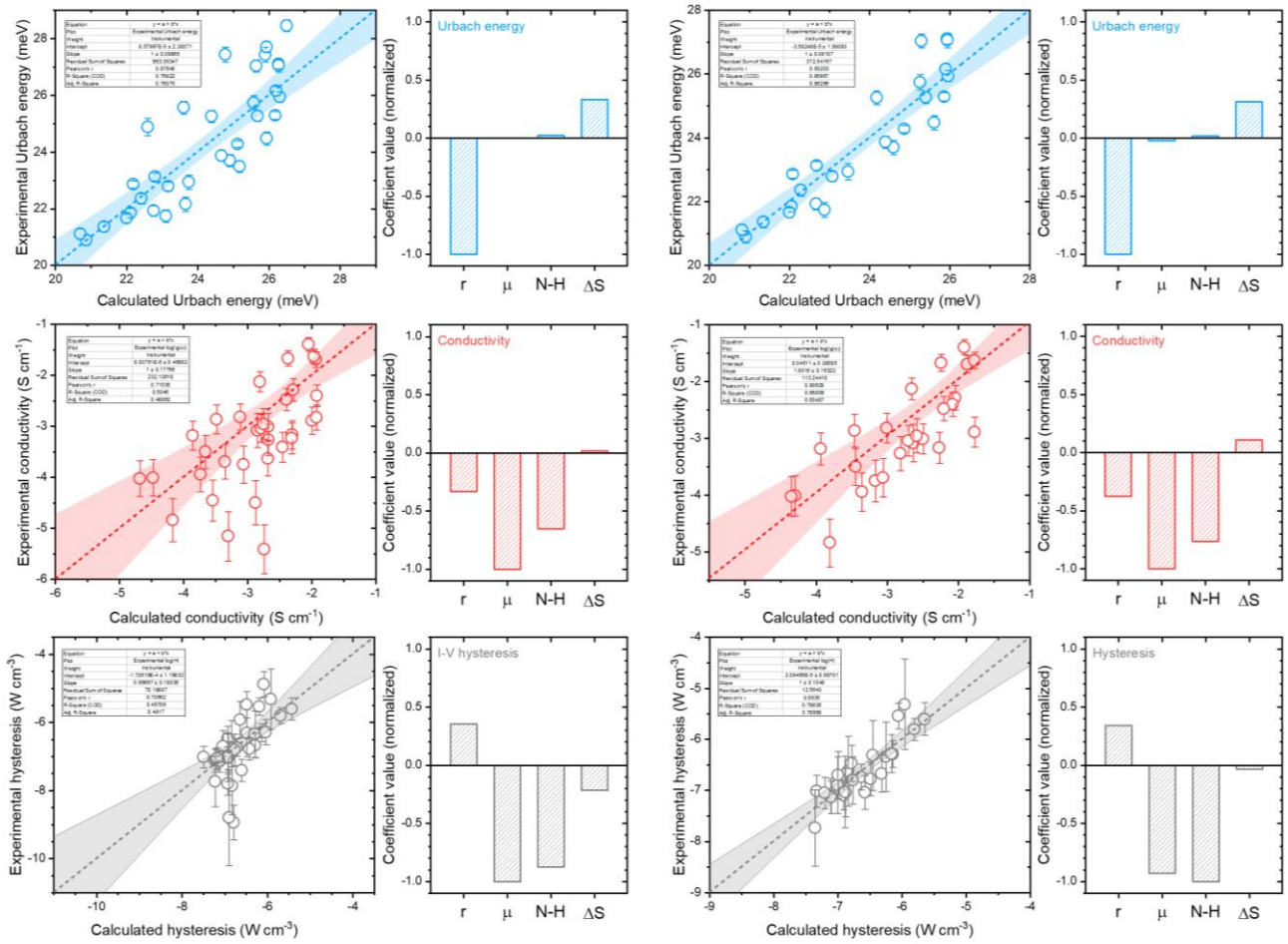


Table S3 - Absolute coefficients' values and errors (using z-scored entries) obtained through multivariate linear regression of data before and after data exclusion, and statistical parameters for a significance level of 0.95 (95%).

Property	Coeff.	33 measurements				25 measurements			
		value	error	t-value	prob > |t|	value	error	t-value	prob > |t|
T → C transition temperature (K)	Intercept	273.7576	2.6316	104.0276	0	275.0550	1.6786	163.8650	0
	r	-16.2296	7.4662	-2.1737	0.0383	-12.1841	5.4227	-2.2469	0.0361
	μ	-1.3373	6.7035	-0.1995	0.8433	9.4047	4.5394	2.0718	0.0514
	NH	21.5267	7.3419	2.9320	0.0066	31.9014	4.6409	6.8739	$< 10^{-4}$
	ΔS	-25.2456	6.4938	-3.8876	0.0006	-33.0192	4.2159	-7.8322	$< 10^{-4}$
Band gap energy (eV)	Intercept	1.4722	0.0012	1223.3303	0	1.4720	0.0009	1680.8613	0
	r	0.0046	0.0031	1.4746	0.1515	0.0044	0.0021	2.0762	0.0510
	μ	0.0303	0.0030	10.1579	$< 10^{-4}$	0.0300	0.0022	13.5596	$< 10^{-4}$
	NH	0.0175	0.0034	5.2382	$< 10^{-4}$	0.0156	0.0027	5.8627	$< 10^{-4}$
	ΔS	-0.0085	0.0028	-3.0760	0.0047	-0.0057	0.0022	-2.6498	0.0154
Urbach energy (meV)	Intercept	24.1995	0.1971	122.7955	0	23.9854	0.1702	140.9257	0
	r	-2.5349	0.5596	-4.5295	0.0001	-2.3657	0.4476	-5.2855	$< 10^{-4}$
	μ	0.0023	0.4835	0.0047	0.9963	-0.0561	0.4009	-0.1399	0.8902
	NH	0.0548	0.5308	0.1033	0.9185	0.0422	0.4232	0.0997	0.9216
	ΔS	0.8458	0.4819	1.7550	0.0902	0.7377	0.4017	1.8365	0.0812
Conductivity (log S cm^{-1})	Intercept	-2.8838	0.1391	-20.7353	0	-2.7404	0.1232	-22.2379	0
	r	-0.5176	0.3770	-1.3729	0.1807	-0.5819	0.3209	-1.8136	0.0834
	μ	-1.5607	0.3381	-4.6157	0.0001	-1.5446	0.2868	-5.3855	$< 10^{-4}$
	NH	-1.0169	0.3808	-2.6705	0.0125	-1.1794	0.3290	-3.5846	0.0017
	ΔS	0.0230	0.2914	0.0788	0.9377	0.1650	0.2535	0.6511	0.5217
I-V hysteresis (log W cm^{-3})	Intercept	-6.6512	0.1089	-61.0806	0	-6.6044	0.0639	-103.4338	0
	r	0.3237	0.2802	1.1555	0.2577	0.3152	0.1447	2.1788	0.0415
	μ	-0.9099	0.2566	-3.5461	0.0014	-0.8520	0.1349	-6.3143	$< 10^{-4}$
	NH	-0.7963	0.2982	-2.6702	0.0125	-0.9178	0.1649	-5.5658	$< 10^{-4}$
	ΔS	-0.1971	0.2628	-0.7503	0.4594	-0.0293	0.1535	-0.1911	0.8504

Supplementary Note 5: Normalized coefficients and lm.beta values

Unlike classical multiple linear regression, which requires strict assumptions about normality, homoscedasticity, independence of residuals, and absence of multicollinearity, the Partial Least Squares (PLS) method is particularly suitable for scenarios with strongly correlated predictors and a small number of samples—exactly the case for the molecular descriptors used in this work. Thus, by normalizing the data using z-scores, the resulting coefficients from the PLS model can be interpreted directly, without the need to fully meet the traditional assumptions of classical linear regression.

In regression analysis with multiple independent variables, the crude coefficients do not allow for an adequate comparison of the relative influence of each descriptor, because each variable has different units and scales. For this reason, the assessment of the relative importance of the predictors should be done using the standardized coefficients, obtained after normalizing the independent and dependent variables. This standardization places all variables on the same scale (mean = 0; standard deviation = 1), allowing the modulus of each standardized coefficient to represent the strength of the effect exerted by that descriptor on the property studied. Conceptually, these coefficients correspond to what, in the R software, is returned by the "lm.beta()" function, widely used to quantify the relative importance of predictors in multiple regression models. Thus, the standardized coefficients used in this work are equivalent to the coefficients returned by lm.beta, as they express the impact of each descriptor after z-score standardization, allowing the relative contribution of each independent variable to the response variable to be evaluated. The normalized parameters and lm.beta values shown in Figure S10 explicitly show these correlations.

In the context of the present study, the standardized coefficients extracted from the PLS model play exactly this role: they indicate, in a quantitative and comparable way, which descriptors of cation A exert greater or lesser influence on the properties evaluated (tetragonal to cubic transition temperature, band gap, Urbach energy, electrical conductivity and hysteresis). Thus, the analysis of these coefficients provides a clear, objective, and physically interpretable measure of the statistical relevance of each described for the engineering of halide perovskite compositions.

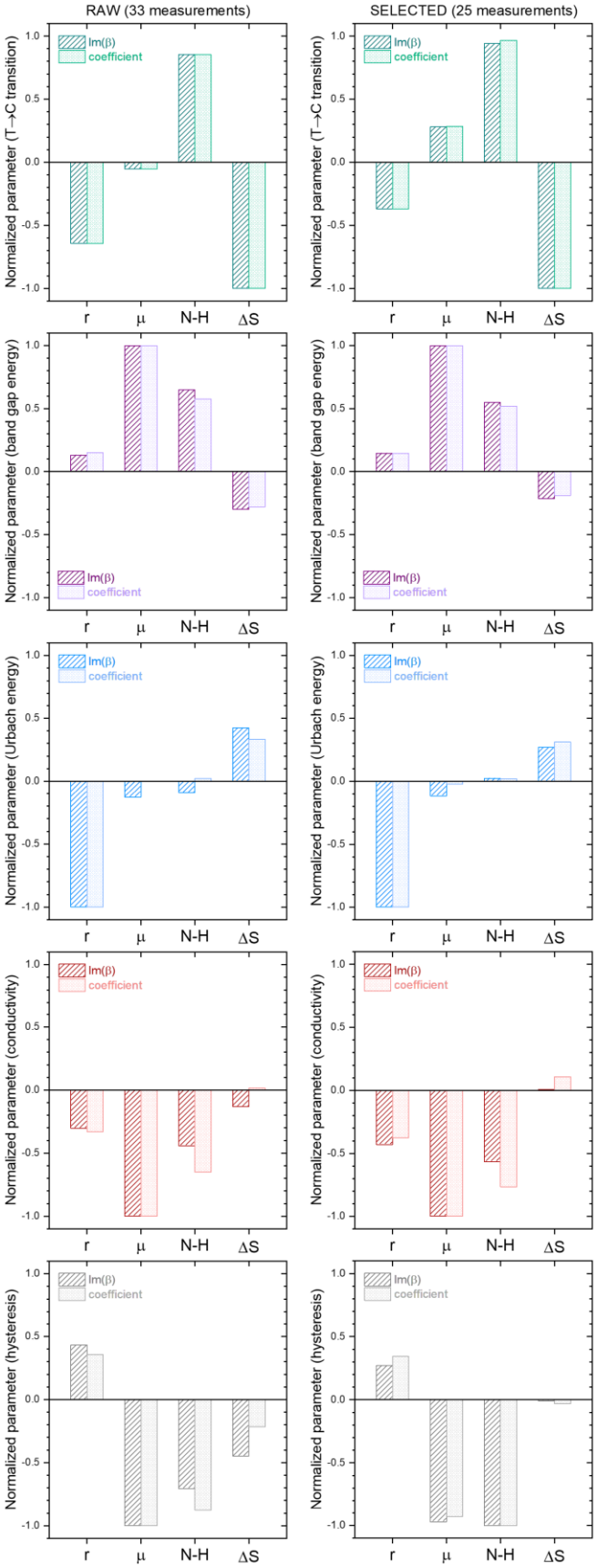


Figure S10 - Comparison of normalized coefficients derived from the multivariate regression and lm. beta values before and after data selection.

Supplementary Note 6: Properties in tetragonal and cubic phases

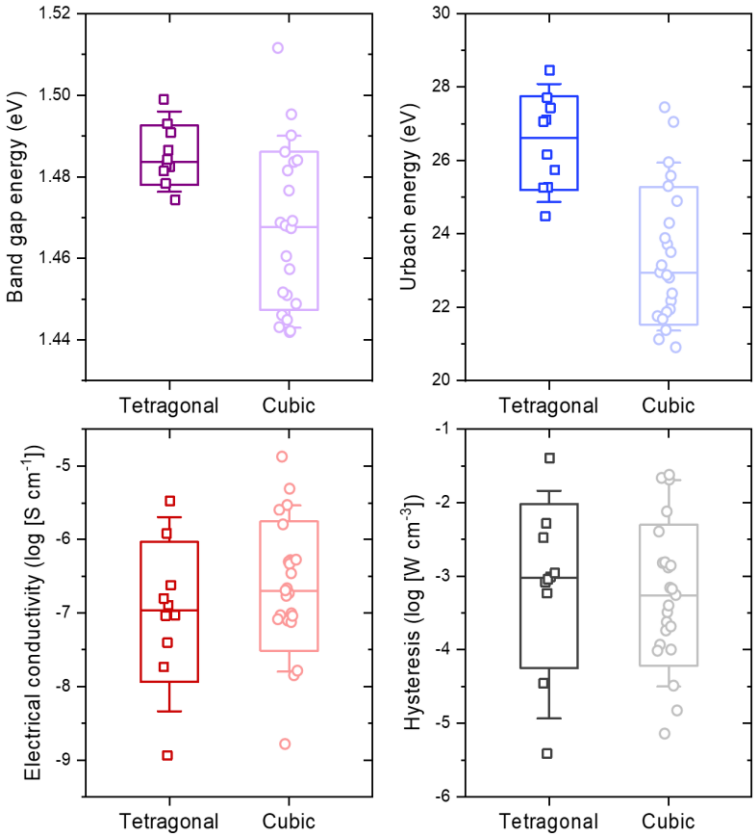


Figure S11 - Comparison between experimental properties for the compositions of tetragonal and cubic phases at room temperature. All 33 points are plotted. Central lines are the median, boxed the standard deviations and whiskers the 10 - 90% percentiles.

Supplementary Note 7: Qualitative predictions of hypothetical systems

To visualize how the results can be used for some qualitative insights on property variations of the two hypothetical $FMA_xMA_{1-x}PbI_3$ and $HMA_xMA_{1-x}PbI_3$ systems, we first estimate FMA^+ and HMA^+ properties. Using a simplified version of the protocol in [6], hydroximethylammonium radius can be estimated by r_A = 146 (ammonium) + 71 (methyl) + 70 (hydroxil) = 286 pm. The Shannon radius of OH^- is 132 pm and of F^- is 128.5 pm for a coordination number of II [7]. Since this difference is very small and irrelevant for a qualitative analysis, the r_A of OH^- and I^- was considered to be approximately equal, that is, 70 pm. Thus, the radius of both organic cations was taken to be of the order of 286 pm. By similarity of radii, the comparison is made between the systems $FMA_xMA_{1-x}PbI_3$, $HMA_xMA_{1-x}PbI_3$, and $GA_xMA_{1-x}PbI_3$, this last one well-known. The schematic variations are shown in Figure S12.

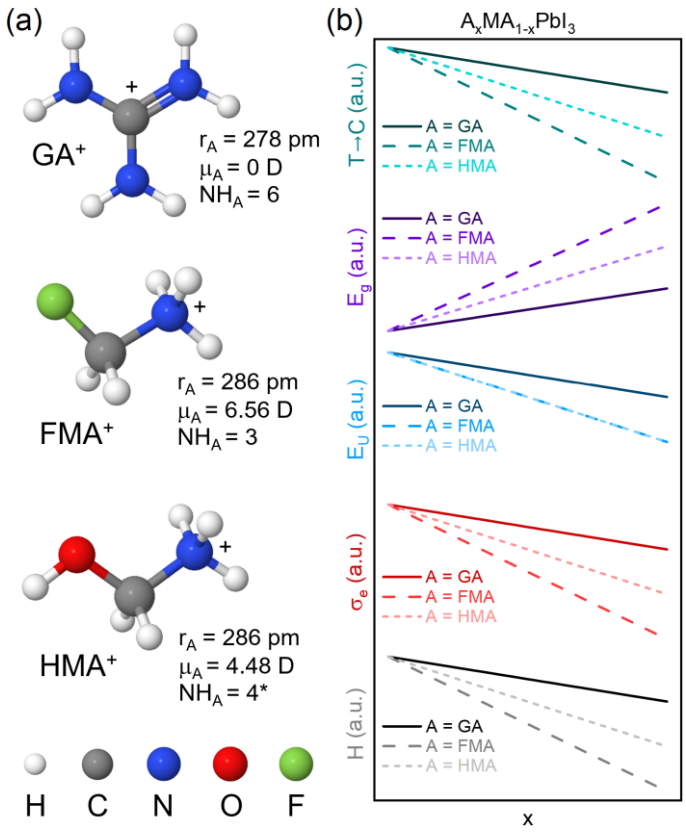


Figure S12 - a) Structure view of the substituting cations in $FMA_xMA_{1-x}PbI_3$, $HMA_xMA_{1-x}PbI_3$, and $GA_xMA_{1-x}PbI_3$ and their properties. The dipole moments were estimated using MolCalc [8]. The number of N-H bonds in HMA^+ was considered to be of four due to the hydrogen atom bound to the oxygen. b) Schematic variation of the studied properties between the three systems with increasing substitution contents. Noteworthy, we are neglecting other type of interactions, such as halogen-halogen bonding [9].

Supplementary Note 8: Comparison of properties of literature

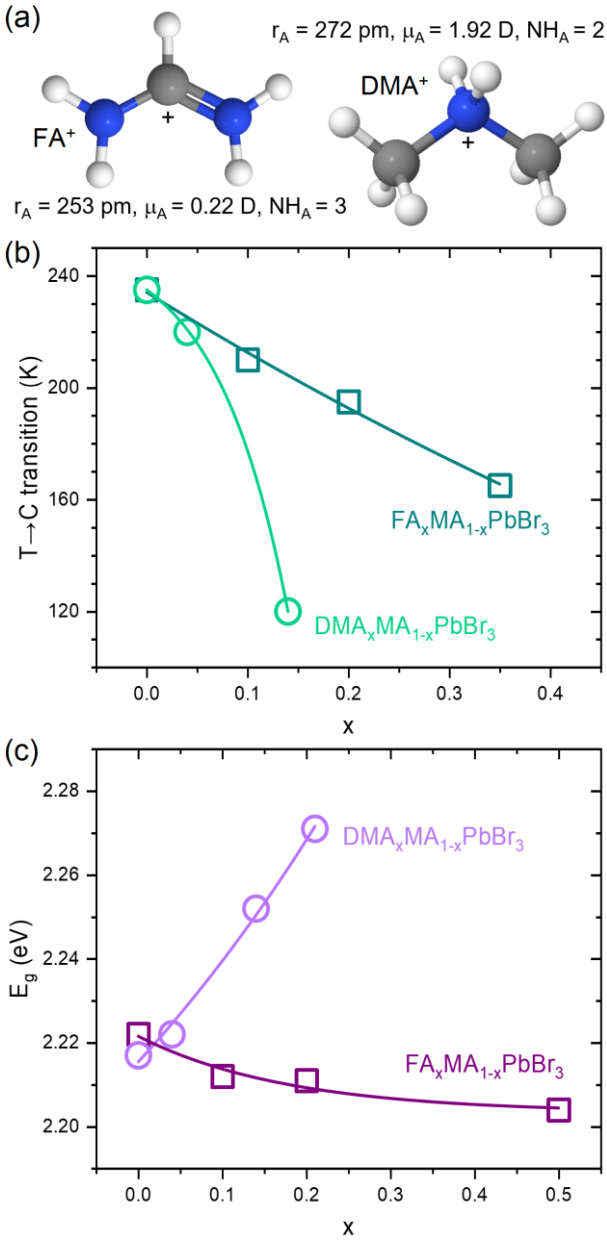


Figure S13 - a) Structure view of the substituting cations in $FA_xMA_{1-x}PbBr_3$ and $DMA_xMA_{1-x}PbBr_3$. Radii are found in [10]. The dipole moments were estimated using MolCalc [8]. b) Tetragonal-to-cubic transition temperature and band gap energy of the two systems as a function of the substitution content. $FA_xMA_{1-x}PbBr_3$ and $DMA_xMA_{1-x}PbBr_3$ data were taken from [11] and [12], respectively.

References

[1] P. Makuła, M. Pacia and W. Macyk, How To Correctly Determine the Band Gap Energy of Modified Semiconductor Photocatalysts Based on UV–Vis Spectra, *J. Phys. Chem. Lett.*, 2018, **9**, 6814–6817.

[2] S. Landi, I. R. Segundo, E. Freitas, M. Vasilevskiy, J. Carneiro and C. J. Tavares, Use and misuse of the Kubelka-Munk function to obtain the band gap energy from diffuse reflectance measurements, *Solid State Communications*, 2022, **341**, 114573.

[3] V. Sarritzu, N. Sestu, D. Marongiu, X. Chang, Q. Wang, S. Masi, S. Colella, A. Rizzo, A. Gocalinska, E. Pelucchi, M. L. Mercuri, F. Quochi, M. Saba, A. Mura and G. Bongiovanni, Direct or Indirect Bandgap in Hybrid Lead Halide Perovskites?, *Advanced Optical Materials*, 2018, **6**, 1701254.

[4] C. Witt, K. Schötz, A. Köhler and F. Panzer, Understanding Method-Dependent Differences in Urbach Energies in Halide Perovskites, *J. Phys. Chem. C*, 2024, **128**, 6336–6345.

[5] R. Kumar, B. Rakheja, N. Lamminen, F. Fasulo, M. A. T. Cachafeiro, C. Hanmandlu, G. K. Grandhi, M. Bag, A. B. Muñoz-García, G. Boschloo, W. Tress, M. Pavone, P. Vivo and E. M. J. Johansson, Mechanistic Insights into Ionic Conduction in Lead Halide Perovskites and Perovskite-Inspired Materials, *Advanced Energy Materials*, 2025, e03331.

[6] G. Kieslich, S. Sun and A. K. Cheetham, Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog, *Chem. Sci.*, 2014, **5**, 4712–4715.

[7] Shannon Radii, http://abulafia.mt.ic.ac.uk/shannon/ptable.php, (accessed 22 October 2025).

[8] J. H. Jensen and J. C. Kromann, The Molecule Calculator: A Web Application for Fast Quantum Mechanics-Based Estimation of Molecular Properties, *J. Chem. Educ.*, 2013, **90**, 1093–1095.

[9] M. L. Ball, J. V. Milić and Y.-L. Loo, The Emerging Role of Halogen Bonding in Hybrid Perovskite Photovoltaics, *Chem. Mater.*, 2022, **34**, 2495–2502.

[10] M. Simenas, A. Gagor, J. Banys and M. Maczka, Phase Transitions and Dynamics in Mixed Three- and Low-Dimensional Lead Halide Perovskites, *Chem. Rev.*, 2024, **124**, 2281–2326.

[11] M. Šimėnas, S. Balčiūnas, Š. Svirskas, M. Kinka, M. Ptak, V. Kalendra, A. Gągor, D. Szewczyk, A. Sieradzki, R. Grigalaitis, A. Walsh, M. Mączka and J. Banys, Phase Diagram and Cation Dynamics of Mixed $MA_{1-x}FA_xPbBr_3$ Hybrid Perovskites, *Chem. Mater.*, 2021, **33**, 5926–5934.

[12] M. Simenas, S. Balciunas, J. N. Wilson, S. Svirskas, M. Kinka, A. Garbaras, V. Kalendra, A. Gagor, D. Szewczyk, A. Sieradzki, M. Maczka, V. Samulionis, A. Walsh, R. Grigalaitis and J. Banys, Suppression of phase transitions and glass phase signatures in mixed cation halide perovskites, *Nat Commun*, 2020, **11**, 5103.

Empirical equations to estimate tetragonal to cubic phase transitions in lead iodide perovskites

Fernando B. Minussi[1,2,3*], Rogerio M. Silva Jr.[4], Thays A. Abreu[5], Jose A. Eiras[3], Eudes B. Araujo[2]

[1] Department of Mechanical Engineering, Santa Catarina State University, Joinville, SC, Brazil

[2] Department of Physics and Chemistry, São Paulo State University, Ilha Solteira, SP, Brazil

[3] Department of Physics, Federal University of São Carlos, São Carlos, SP, Brazil

[4] Department of Electrical Engineering, São Paulo State University, Ilha Solteira, SP, Brazil

[5] Department of Mathematics, São Paulo State University, Ilha Solteira, SP, Brazil

Abstract

Phase transitions in halide perovskites are well-studied due to their fundamental and applied importance. A wide range of different symmetries may appear depending on the composition, temperature, etc. For lead iodide perovskites, fortunately, the transition near room temperature is almost always between tetragonal and cubic phases. Knowing these transitions and the stable phase under a given condition can have technological implications. Although extensively studied, the current literature lacks simple tools to estimate these transition temperatures. To address this, we meticulously collected literature and compiled a database containing phase information for a wide range of $APbI_3$-type compositions. Data from single-phase systems were converted into easily obtainable physicochemical descriptors, and regression tools were applied, namely multivariate linear analysis and the recent method known as SISSO (sure independence screening and sparsifying operator). Simple empirical equations were determined and applied to specific systems, and their limitations and potentialities are discussed. This work advances the understanding of the topic and provides additional insights of interest for future developments in the field.

Keywords: Halide perovskites, A-site cations, physicochemical descriptors, phase transitions, multivariate linear regression, SISSO, limit of solubility.

* fbminussi@gmail.com / f.minussi@unesp.br

Halide perovskites (HPs) are a well-known class of materials with broad optoelectronic and electrochemical applications [1,2,3]. It is greatly due to the large number of different compounds and solid solutions that adopt the perovskite structure. They have a stoichiometry based on ABX_3, in which, most often, A is an organic or inorganic monovalent cation (such as Cs^+, methylammonium, MA^+, formamidinium, FA^+), B a divalent cation (Pb^{2+}, Sn^{2+}, Ge^{2+}), and X a halide anion (Cl^-, Br^-, I^-) [4,5], where the tridimensional (3D) framework is formed by corner-sharing connection of BX_6 octahedra [6,7].

When considering the entire class of HPs, phases belonging to all seven crystal systems are found in several symmetries (i.e., space groups) over different temperature ranges [8,9]. Even when the B and X sites are held fixed, the present phases and transitions are strongly A-site dependent. The relevance of polymorphism and phase transitions in halide perovskites was the topic of dedicated recent reviews [5,10]. Focusing on the A-site, many factors, including cation sizes, rotation, ordering, and orientation, as well as the possible formation of hydrogen bonds and their strength [11], affect what the stable phases and the temperatures at which phase transitions may occur.

For almost all lead iodide-based perovskites, structural transitions that occur near room temperature are between the tetragonal and cubic phases. From an applied perspective, understanding phase transitions is important to minimize the risk of malfunction or failure due to dimensional changes in devices operating over a range of temperatures. Even though it leads to minor volume changes, the tetragonal-cubic phase transition in perovskites adversely affects charge transport and film stability [12]. In particular, tracking stable phases and transitions is important since there is a general preference for cubic halide perovskites in solar cells [13]. Indeed, cubic halide perovskites tend to have lower hysteresis, band gaps, and Urbach energies, and higher electric conductivities, characteristics commonly desired in devices [14]. On the other hand, it is claimed that ferroic polarization in hybrid perovskites is crucial for enhancing photovoltaic performance and developing potential electronic applications [15,16]. Without entering the long debate of ferroelectricity in $MAPbI_3$-based materials [17,18,19,20], it is a fact that properties like piezo and ferroelectricity are not features of centrosymmetric cubic phases. Having in mind the several interesting applications of ferroic halide perovskites [21,22,23], it turns out that some advantages may arise in the tetragonal phase.

Intriguingly, the current literature lacks simple tools or models capable of predicting or estimating phase transition temperatures in HPs, even in lead iodide perovskites with variations only at the A-site. To overcome this barrier, we here propose two simple models based on

parameters that can be easily determined for a predefined composition. The equations were determined empirically from literature available data using two approaches, namely, a multivariate linear analysis (MLA), analogously to recently employed [14], and the sure independence screening and sparsifying operator (SISSO) [24,25], a recently developed machine learning method for identifying physical models of a given property from initial proposed features or descriptors. Briefly, we constructed a database from a meticulous analysis of the phase data in the literature for perovskites of the type $APbI_3$, which is available to readers in Supplementary Note 1 and the associated link. To simplify the entire chemical space, these A-site varying compositions were converted to four physicochemical descriptors, namely, the effective radius (r_{eff}), dipole moment (μ_{eff}), number of N-H bonds (n_{NH}), and the configurational entropy of mixing (ΔS_{mix}). The equations used for these conversions are given in Table 1. Beyond their simplicity and prompt determination, these descriptors are characteristics of the A-site cations that should alter the octahedral tilting and distortions, bond angles and lengths etc. of the inorganic framework, leading to different phase transition temperatures, but are much easier to determine. The methodology for model construction is described in Supplementary Note 2.

It is important to note the reasons we chose lead iodide perovskites as a representative for the present study. Iodine perovskites have the lowest band gaps when combined with Sn^{2+}, but with Pb^{2+}, the band gaps are only 0.2-0.3 eV larger [26]. On the other hand, Sn has additional degradation pathways, in particular, oxidation of Sn^{2+} to Sn^{4+}, i.e., Pb-based perovskites tend to be more stable [27]. Moreover, bromide and chloride-containing perovskites have larger band gaps [28,29]. Also, using other B and X sites leads to different types of transitions, increasing the complexity beyond what is intended for a first approximation as proposed in this work. Lastly, there is more information available on the phase transitions of lead iodide perovskites than on those of other types of halide perovskites. That said, the results are shown in Figure 1.

In Figure 1a, we show the correlation between the experimental and calculated values using the MLA approach. A reasonable R^2 value is shown using the derived equation above the plot. Still, we refined the correlation, removing only eight out of 125 initial values based on the criterion $\left|T_{\mathrm{calc}} - T_{\mathrm{exp}}\right|/T_{\mathrm{exp}} \geq 0.1$. The results shown in Figure 1b demonstrate a significant improvement after minor data polishing. In Figure 1c, we show the lm. beta values, as discussed recently [14], that can be used to infer the magnitude and sense of the descriptors on the calculated property. Specifically, it is clear that the radius and entropy strongly reduce the tetragonal-to-cubic phase transition temperature. On the other hand, the number of N-H bonds

has a strong positive effect, increasing temperatures, while the dipole has a weaker effect than the other descriptors. This is, with minor relative magnitude differences, the same trend we recently observed [14], corroborating the approach. In Figure 1d, we show the results from the SISSO approach, yielding a good R^2 value. The three graphs show the same plot but in different visualizations. A crucial distinction between the MLA- and SISSO-derived models is that the MLA cannot describe results far from the high-concentration region, i.e., below 200 K or above 350 K. This is best illustrated in the graph given in the Supplementary Note 3. These data could not be included in the MLA model because they lower the R^2, and the calculated errors at the extremes are already immense. When excluded, the R^2 are better, but extrapolated estimates at the extremes are still poor. On the other hand, the SISSO approach describes these values well. However, by doing so, the best SISSO model do not account for the entropy as a relevant parameter, contrary to MLA. Also, as shown in Figure 1 and Supplementary Note 3, the results from 200 to 350 K are better using MLA.

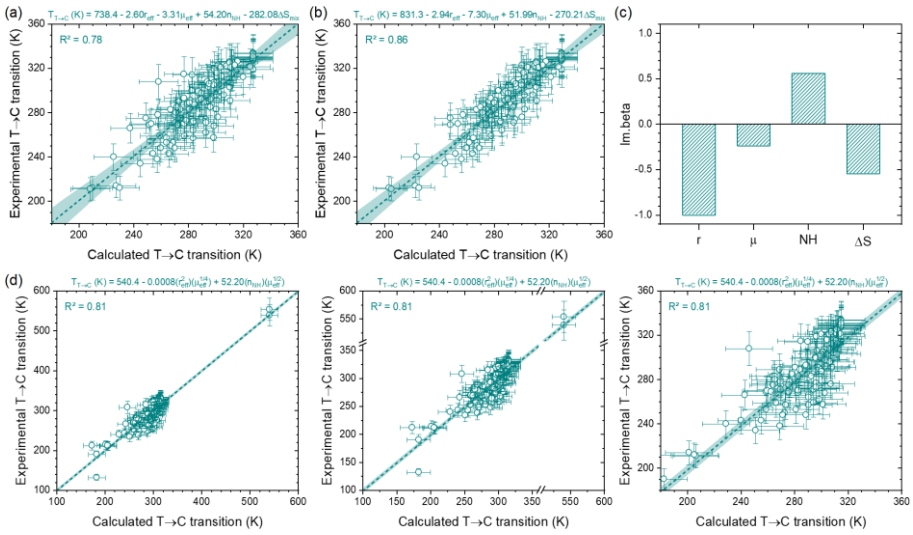


Figure 1 - a) Preliminary and b) refined experimental versus calculated tetragonal to cubic phase transitions derived from the MLA regression. c) lm. beta of the proposed A-site descriptors. d) SISSO-derived results visualized in different regions. In all graphs, error bars on the x-axis are RMSE values, and on the y-axis, an arbitrarily employed 5% error. Dashed lines are x = y, and shaded areas are the 99% confidence bands.

Now, some aspects need to be highlighted. In our analyses, we did not discriminate between transition temperatures obtained using different techniques, nor between those measured during cooling or heating. Technique-dependence and thermal hysteresis may be large, and they should constitute significant sources of errors included in the RMSE values. Also, we recently discussed qualitative impressions on which parameters may be relevant/useful to explain why the different A-site compositions in lead iodide perovskites

provoke the experimentally observed effects in phase transitions. Here, we attempted to do it analogously, but in a quantitative manner. While the models may be viewed as too simplistic, we believe they are reasonably powerful, as they can provide guidance within minutes on the phase transitions of *a priori* defined tridimensional lead iodide perovskite composition. Doing so using *ab initio* computational methods might be more accurate, but covering a large A-site chemical space is a demanding task if rough estimations are sufficient. On the other hand, machine learning tools could be useful and less computationally intensive, but most would not yield what we truly want here: a simple equation. Lastly, relative energetics and dynamics between phases may vary with the different specific symmetries of phases, a contrast not considered here. Instead, we only assign tetragonal and cubic phases, disregarding the diversity of space groups. For example, it doesn't matter here if the tetragonal phase is $I4cm$ or $I4/mcm$.

Table 1 - Summary of the A-site variables to be determined for a given composition and derived empirical equations from literature data using multivariate linear analysis (MLA) and sure independence screening and sparsifying operator (SISSO) approaches. x_i is the molar fraction of each A-site cation of properties r_{A}, μ_{A}, and $\mathrm{NH_A}$, respectively, the radius, dipole moment, and number of N-H bonds, as given in Supplementary Note 1. R is the gas constant. RMSE is the root mean squared error.

A-SITE variables	$r_{\mathrm{eff}} = \sum_{i=1}^{\infty} x_i(r_{\mathrm{A},i})$	$\mu_{\mathrm{eff}} = \sum_{i=1}^{\infty} x_i(\mu_{\mathrm{A},i})$	$n_{\mathrm{NH}} = \sum_{i=1}^{\infty} x_i(\mathrm{NH}_{\mathrm{A},i})$	$\Delta S_{\mathrm{mix}} = -\frac{R}{5}\sum_{i=1}^{\infty} x_i(\ln x_i)$
MLA	$T_{\mathrm{T}\to\mathrm{C}}\ (\mathrm{K}) = a + b(r_{\mathrm{eff}}) + c(\mu_{\mathrm{eff}}) + d(n_{\mathrm{NH}}) + e(\Delta S_{\mathrm{mix}})$ a = 831.3 K, $b = -2.94$ K pm^{-1}, $c = -7.30$ K D^{-1}, d = 51.99 K, $e = -270.21$ K R^{-1}, RMSE = 11.4 K			
SISSO	$T_{\mathrm{T}\to\mathrm{C}}\ (\mathrm{K}) = a + b(r_{\mathrm{eff}}^2)(\mu_{\mathrm{eff}}^{1/4}) + c(n_{\mathrm{NH}})(\mu_{\mathrm{eff}}^{1/2})$ a = 540.4 K, $b = -8.10^{-4}$ K pm^{-2} D$^{-1/4}$, c = 52.20 K D$^{-1/2}$, RMSE = 17.6 K			

With all that said, the summary of results is shown in Table 1. Our impressions on the results lead to the following way to apply the models: step 1) set the desired A-site composition; step 2) obtain, for each cation, the radius from the methodology of [4], the dipole moments from [30], and the number of N-H bonds as the sum of moieties capable of forming hydrogen bonds, such as N-H, O-H, and S-H; Step 3) convert the A-site composition to the variables r_{eff}, μ_{eff}, n_{NH}, and ΔS_{mix}; step 4) Insert the values in the MLA and SISSO equations. MLA seems to work better in general, but is highly not recommended when the perovskites are Cs^+- and AZR^+-based. These cations lead to extreme transition temperatures, with AZR^+ at the lowest and Cs^+ at the highest. Until additional transition temperatures in other systems are available, clear preferences between MLA and SISSO cannot be suggested.

To obtain additional insights into the applicability of the equations, we analyzed several individual cases, shown in Figure 2. Both models yield similar results in both systems, with slightly improved accuracy with MLA (as shown in Figure 2a). On the other hand, as shown in Figure 2b, when Cs^+ is used, the MLA model is not appropriate, while SISSO may eventually work. In Figure 2c, it can be seen that the models struggle to describe "symmetrization plot twists" [31], i.e., the inversion at the reduction trend of the transition temperatures. However, it is interesting that the models can describe the systems to some extent. Also, intriguingly, both models apparently work very well for the $en_xMA_{1-x}Pb_{1-0.7x}I_{3-0.4x}$ for increasing x up to the "plot twist". Worthy of note, these compositions were the only ones in Figure 2 that were not used to create the models. This is curious since, unlike all compositions added to the database, the en^{2+} cation is divalent, leading to stoichiometric deviations at the Pb and I sites. In this case, the configurational entropy of mixing has to be calculated considering all the sublattices, thus $\Delta S_{\text{mix}} = -R \sum_S a^S \sum_i y_i^S \ln y_i^S$, a^S is the fraction of each sublattice (S) in the stoichiometric formula and y_i^S is the mol fraction of each constituent species (i) in its sublattice [32].

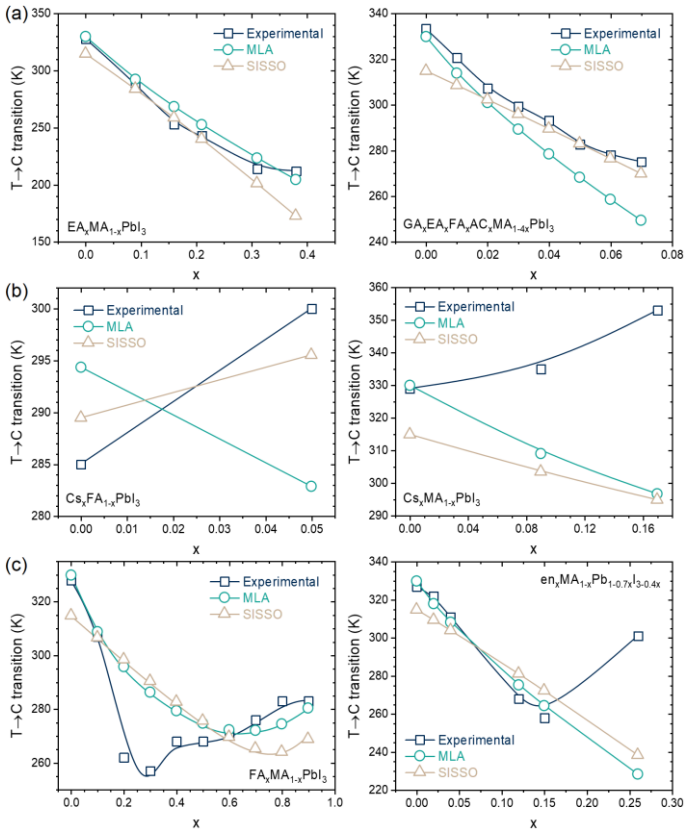


Figure 2 - Accuracy evaluation of the equations in some systems with known tetragonal to cubic phase transition temperatures. a) Two cases with straightforward applicability (left [33] and right [34]). b) Cs^+-containing compositions, evidencing the limitations (left [35] and right [36]). c) Systems with a trending inversion (left [37] and right [38]). Lines are guides for the eyes.

Finally, we tested the models by predicting transition temperatures in systems for which transition temperatures have not been experimentally determined, but for which some phase data are available. The results are summarized in Figure 3, such that the x dependent transition temperature of the $IM_xMA_{1-x}PbI_3$ is shown in Figure 3a. This system has a solubility limit of around $x \sim 0.10$, and all solid solutions are tetragonal at room temperature [39]. If we take 295 K as a reference, then all compositions must have a transition temperature above room temperature. This is the case as determined by the MLA equation, and is closely matched by the SISSO equation. As for the $DMA_xMA_{1-x}PbI_3$ system, the solubility limit is of $x \sim 0.15$, and solid solutions are tetragonal up to $x \sim 0.050$, and cubic at $x \sim 0.075$ or higher [39]. As shown in Figure 3b, both equations perform well, with SISSO being a bit better. For $TMS_xFA_{1-x}PbI_3$ compositions, shown in Figure 3c, the solubility limit is only $x \sim 0.05$ and are cubic [40], well predicted with both equations. Finally, for $en_xFA_{1-x}Pb_{1-0.7x}I_{3-0.4x}$, compositions are cubic up to $x \sim 0.27$ and tetragonal close to the solubility limit at $x \sim 0.39$ [41]. This behavior indicates a trend inversion in the transition temperature, which is not predicted in Figure 3d with any of the equations. Since the MLA curve decays too steeply, it is possible that it is farther from what would be observed experimentally.

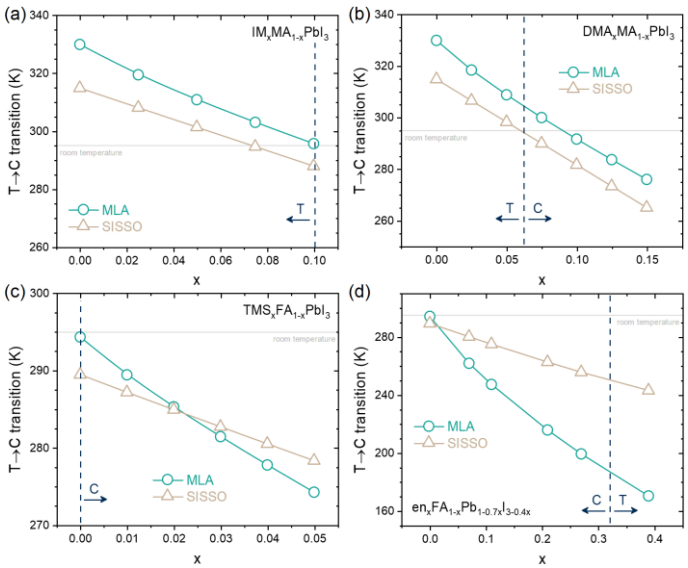


Figure 3 - MLA and SISSO testing of systems with limited phase information: a) $IM_xMA_{1-x}PbI_3$, b) $DMA_xMA_{1-x}PbI_3$, c) $TMS_xFA_{1-x}PbI_3$, and d) $en_xFA_{1-x}Pb_{1-0.7x}I_{3-0.4x}$. The horizontal gray line represents room temperature. The arrows indicate the tetragonal (T) and cubic (C) phase fields as x varies. Lines are guides for the eyes.

As a relevant topic, after mentioning the solubility limits in some HPs, we conclude the discussion with a few words on the formability of single-phase solid solutions. In $MAPbI_3$-

based perovskites, the tendency is for solubility limits to increase when the substituting cations have close, but larger, radii and more N-H bonds than MA^+ [42]. However, the same principles do not apply to $FAPbI_3$-based ones. For example, in $Cs_xFA_{1-x}PbI_3$ and $MA_xFA_{1-x}PbI_3$, both Cs^+ and MA^+ are smaller than FA^+, but the first one has a solubility limit of only $x \sim 0.10$ [43] while the second is a complete solid solution [44]. Changing to the $DMA_xFA_{1-x}PbI_3$ system, the DMA^+ cation is larger than FA^+ and does not even form 3D perovskites regardless of the $x > 0$ value [45]. On the other hand, as discussed, the $en_xFA_{1-x}Pb_{1-0.7x}I_{3-0.4x}$ system does form 3D perovskites with a quite large solubility limit, but en^{2+} is even larger [41] than DMA^+. Having outlined this scenario, it is clear that a general understanding of the features that guide phase formability and solid solutions in HPs has not yet been achieved, despite the immense body of information on these materials. While some attempts were proposed [46,47,48], a generalization on the formability of 3D halide perovskites, and their single-phase solid solutions, remains an unresolved issue. Inspired by the work in [46], we employed principal component analysis, clustering, and machine learning methods, including SISSO, trying to find ways (parameters, combination of parameters, operators, etc.) to separate and discriminate the compounds of the database into groups, namely, non-perovskites and perovskites, the latter subdivided into single- or mixed-phase. However, none of our approaches was successful. We humbly invite the community to explore this challenging and important topic.

In conclusion, we derived two simple empirical equations to estimate the tetragonal-to-cubic phase transition temperatures in lead iodide perovskites. The models are applicable to varying degrees depending on system-specific characteristics. Suboptimal cases include systems containing Cs^+ or AZR^+, and systems that exhibit trending inversions upon varying substitution contents. Still, we provide simple ways to estimate this important characteristic and discuss new insights into which parameters control the relative phase stability in HPs. We also believe that the provided database in this work is an important complement to the recent reviews in phase transitions of HPs and summarizes some general literature regarding phase data in lead iodide perovskites to facilitate future endeavors for the interested readership.

See the supplementary material for details on the properties of several A-site cations, phase information collected from literature, regression development, and comparison of results from MLA and SISSO.

This work has been supported by Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-INCT: 406322/2022-8 and CNPq Grant: 315586/2023-0) and

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-DAAD: 88881.895061/2023-01). The author F.B. Minussi would like to thank the CNPq for their Postdoctoral Fellowship (Grant: 157815/2025-0).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose

Author Contributions

F. B. Minussi: conceptualization (leading), methodology (leading), validation (leading), formal analysis (leading), investigation (supporting), writing - original draft (leading), writing - review and editing (leading), and visualization (leading). R. M. Silva Jr.: conceptualization (support), methodology (support), software (leading), validation (support), formal analysis (leading), investigation (leading), writing - original draft (support), writing - review and editing (support), and visualization (support). T. A. Abreu: conceptualization (support), methodology (support), software (leading), validation (support), formal analysis (support), writing - original draft (support), writing - review and editing (support). J. A. Eiras: conceptualization (support), methodology (support), resources (leading), writing - original draft (support), writing - review and editing (support), supervision (leading), project administration (leading), and funding acquisition (leading). E. B. Araújo: conceptualization (support), methodology (support), resources (leading), writing - original draft (support), writing - review and editing (support), supervision (leading), project administration (leading), and funding acquisition (leading).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

REFERENCES

[1] C. He, and X. Liu, “The rise of halide perovskite semiconductors,” Light Sci Appl **12**(1), 15 (2023).

[2] L. Zhang, J. Miao, J. Li, and Q. Li, “Halide Perovskite Materials for Energy Storage Applications,” Adv Funct Materials **30**(40), 2003653 (2020).

[3] S.I. Seok, and T.-F. Guo, “Halide perovskite materials and devices,” MRS Bulletin **45**(6), 427–430 (2020).

[4] G. Kieslich, S. Sun, and A.K. Cheetham, "Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog," Chem. Sci. **5**(12), 4712–4715 (2014).

[5] M. Simenas, A. Gagor, J. Banys, and M. Maczka, "Phase Transitions and Dynamics in Mixed Three- and Low-Dimensional Lead Halide Perovskites," Chem. Rev. **124**(5), 2281–2326 (2024).

[6] J. Breternitz, and S. Schorr, "What Defines a Perovskite?," Advanced Energy Materials **8**(34), 1802366 (2018).

[7] Q.A. Akkerman, and L. Manna, "What Defines a Halide Perovskite?," ACS Energy Lett. **5**(2), 604–610 (2020).

[8] O.I. Kucheriv, V.Y. Sirenko, and I.A. Gural'skiy, "Small Cations for Engineering 3D Hybrid Perovskites," Chemistry A European J **31**(40), e202500765 (2025).

[9] M.P. Hautzinger, W. Mihalyi-Koch, and S. Jin, "A-Site Cation Chemistry in Halide Perovskites," Chem. Mater. **36**(21), 10408–10420 (2024).

[10] A. Alaei, A. Circelli, Y. Yuan, Y. Yang, and S.S. Lee, "Polymorphism in metal halide perovskites," Mater. Adv. **2**(1), 47–63 (2021).

[11] S. Liu, R. Guo, and F. Xie, "The effects of organic cation rotation in hybrid Organic-Inorganic Perovskites: A critical review," Materials & Design **221**, 110951 (2022).

[12] R. Yuan, K. Chen, and X. Zheng, "Strain–Phase Transition Interplay in Perovskites: Coupling Mechanisms and Regulation Strategy," ACS Energy Lett. **11**(2), 1089–1099 (2026).

[13] Q. Wei, H. Liang, Y. Haruta, M. Saidaminov, Q. Mi, M. Saliba, G. Cui, and Z. Ning, "From tetragonal to cubic: perovskite phase structure evolution for high-performance solar cells," Science Bulletin **68**(2), 141–145 (2023).

[14] F. Minussi, R. Silva Jr., M, Teixeira, R. Bonini, T. Abreu, J. Eiras, and E. Araujo, "A-site composition-property relations in halide perovskites: which descriptors do what?". Under review.

[15] Y. Dou, X. Wang, N.W.G. Smith, P. Behera, R.H. Mudiyanselage, B. Guzelturk, D.A. Walko, Y. Pleimling, S. Liu, N. Nici, C. Slebodnick, B. Dryzhakov, B. Hu, A. Raja, R. Ramesh, G.A. Khodaparast, Y. Yan, and L. Quan, "Long-lived photoinduced polar states in metal halide perovskites," Nat Commun **16**(1), 7230 (2025).

[16] Y. Liu, D. Kim, A.V. Ievlev, S.V. Kalinin, M. Ahmadi, and O.S. Ovchinnikova, "Ferroic Halide Perovskite Optoelectronics," Adv Funct Materials **31**(36), 2102793 (2021).

[17] B. Huang, Z. Liu, C. Wu, Y. Zhang, J. Zhao, X. Wang, and J. Li, "Polar or nonpolar? That is not the question for perovskite solar cells," National Science Review **8**(8), nwab094 (2021).

[18] S. Qin, S. Yi, Y. Xu, Z. Mi, J. Zhao, X. Tian, H. Guo, Y. Jiao, G. Zhang, and J. Lu, "Ferroic alternation in methylammonium lead triiodide perovskite," EcoMat **3**(5), e12131 (2021).

[19] H. Zheng, and K.P. Loh, "Ferroics in Hybrid Organic–Inorganic Perovskites: Fundamentals, Design Strategies, and Implementation," Advanced Materials **36**(14), 2308051 (2024).

[20] F. Ambrosio, F. De Angelis, and A.R. Goñi, "The Ferroelectric–Ferroelastic Debate about Metal Halide Perovskites," J. Phys. Chem. Lett. **13**(33), 7731–7740 (2022).

[21] W. Zheng, X. Wang, X. Zhang, B. Chen, H. Suo, Z. Xing, Y. Wang, H. Wei, J. Chen, Y. Guo, and F. Wang, "Emerging Halide Perovskite Ferroelectrics," Advanced Materials **35**(21), 2205410 (2023).

[22] T. Zhang, H. Luo, M. Abdi-Jalebi, H. Chen, and L. Zuo, “Perovskite solar cells with ferroelectricity,” Information & Functional Materials **1**(1), 87–107 (2024).

[23] J.N. Wilson, J.M. Frost, S.K. Wallace, and A. Walsh, “Dielectric and ferroic properties of metal halide perovskites,” APL Materials **7**(1), 010901 (2019).

[24] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L.M. Ghiringhelli, “SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates,” Phys. Rev. Materials **2**(8), 083802 (2018).

[25] T.A.R. Purcell, M. Scheffler, and L.M. Ghiringhelli, “Recent advances in the SISSO method and their implementation in the SISSO++ code,” The Journal of Chemical Physics **159**(11), 114110 (2023).

[26] F. Yang, and K. Zhu, “Advances in Mixed Tin-Lead Narrow-Bandgap Perovskites for Single-Junction and All-Perovskite Tandem Solar Cells,” Advanced Materials **36**(31), 2314341 (2024).

[27] L. Lanzetta, T. Webb, N. Zibouche, X. Liang, D. Ding, G. Min, R.J.E. Westbrook, B. Gaggio, T.J. Macdonald, M.S. Islam, and S.A. Haque, “Degradation mechanism of hybrid tin-based perovskite solar cells and the critical role of tin (IV) iodide,” Nat Commun **12**(1), 2853 (2021).

[28] S.A. Kulkarni, T. Baikie, P.P. Boix, N. Yantara, N. Mathews, and S. Mhaisalkar, “Band-gap tuning of lead halide perovskites using a sequential deposition process,” J. Mater. Chem. A **2**(24), 9221–9225 (2014).

[29] D. Ory, G. Ammirati, B. Paci, P. Baranek, A. Generosi, O. Fournier, S. Turchini, F. Toschi, J. Guillemoles, J. Barichello, F. Matteocci, A.D. Carlo, V. Milotti, P.M. Sheverdyaeva, P. Moras, D. Catone, and S. Cacovich, “Effect of Chlorine Inclusion in Wide Band Gap $FAPbBr_3$ Perovskites,” Advanced Optical Materials **12**(36), 2401212 (2024).

[30] J.H. Jensen, and J.C. Kromann, “The Molecule Calculator: A Web Application for Fast Quantum Mechanics-Based Estimation of Molecular Properties,” J. Chem. Educ. **90**(8), 1093–1095 (2013).

[31] M. Mączka, M. Ptak, K. Fedoruk, D. Stefańska, A. Gągor, J.K. Zaręba, and A. Sieradzki, “The lattice symmetrization worked, but with a plot twist: effects of methylhydrazinium doping of $MAPbI_3$ on phase transitions, cation dynamics and photoluminescence,” J. Mater. Chem. C **12**(4), 1396–1405 (2024).

[32] S.J. McCormack, and A. Navrotsky, “Thermodynamics of high entropy oxides,” Acta Materialia **202**, 1–21 (2021).

[33] M. Šimėnas, S. Balčiūnas, A. Gągor, A. Pieniążek, K. Tolborg, M. Kinka, V. Klimavicius, Š. Svirskas, V. Kalendra, M. Ptak, D. Szewczyk, A.P. Herman, R. Kudrawiec, A. Sieradzki, R. Grigalaitis, A. Walsh, M. Mączka, and J. Banys, “Mixology of $MA_{1-x}EA_xPbI_3$ Hybrid Perovskites: Phase Transitions, Cation Dynamics, and Photoluminescence,” Chem. Mater. **34**(22), 10104–10112 (2022).

[34] F.B. Minussi, and E.B. Araújo, “On the prospects of high-entropy organic A-site halide perovskites,” Phys. Chem. Chem. Phys. **26**(41), 26479–26488 (2024).

[35] A. Mishra, M.A. Hope, M. Grätzel, and L. Emsley, “A Complete Picture of Cation Dynamics in Hybrid Perovskite Materials from Solid-State NMR Spectroscopy,” J. Am. Chem. Soc. **145**(2), 978–990 (2023).

[36] M. Halder, S. Mal, S. Kalimuddin, S. Balti, H. Acharya, K. Chatterjee, P. Pal, M. Mondal, and S. Das, “Enhancement of structural phase transition temperature of methylammonium lead iodide perovskite by Cesium (Cs) incorporation: Impact on optoelectronic property,” Materials Research Bulletin **196**, 113890 (2026).

[37] O.J. Weber, B. Charles, and M.T. Weller, "Phase behaviour and composition in the formamidinium–methylammonium hybrid lead iodide perovskite solid solution," J. Mater. Chem. A **4**(40), 15375–15382 (2016).

[38] G. Rimkus, S. Balčiūnas, M. Ptak, S. Smółka, V. Klimavičius, D. Szewczyk, A. Sieradzki, V. Kalendra, J. Banys, M. Mączka, and M. Šimėnas, "Phase Diagram and Dielectric Response of Hybrid Lead Halide Hollow Perovskites: A Universal Behavior of Molecular Cation Mixing," Chem. Mater. **36**(15), 7397–7405 (2024).

[39] C.L.C. Ellis, H. Javaid, E.C. Smith, and D. Venkataraman, "Hybrid Perovskites with Larger Organic Cations Reveal Autocatalytic Degradation Kinetics and Increased Stability under Light," Inorg. Chem. **59**(17), 12176–12186 (2020).

[40] S. Sandhu, M.M. Rahman, M. Senthilkumar, B. Yadagiri, J. Park, K. Yoo, and J.-J. Lee, "α-FAPbI3 phase stabilization using aprotic trimethylsulfonium cation for efficient perovskite solar cells," Journal of Power Sources **551**, 232207 (2022).

[41] I. Spanopoulos, W. Ke, C.C. Stoumpos, E.C. Schueller, O.Y. Kontsevoi, R. Seshadri, and M.G. Kanatzidis, "Unraveling the Chemical Nature of the 3D 'Hollow' Hybrid Halide Perovskites," J. Am. Chem. Soc. **140**(17), 5728–5742 (2018).

[42] F. Brondani Minussi, R.M. da Silva, and E. Borges Araújo, "Does it mix? Insights and attempts to predict the formability of single phase mixed A-cation lead iodide perovskites," Inorganic Chemistry Frontiers **10**(21), 6129–6133 (2023).

[43] B. Charles, M.T. Weller, S. Rieger, L.E. Hatcher, P.F. Henry, J. Feldmann, D. Wolverson, and C.C. Wilson, "Phase Behavior and Substitution Limit of Mixed Cesium-Formamidinium Lead Triiodide Perovskites," Chem. Mater. **32**(6), 2282–2291 (2020).

[44] A. Francisco-López, B. Charles, M.I. Alonso, M. Garriga, M. Campoy-Quiles, M.T. Weller, and A.R. Goñi, "Phase Diagram of Methylammonium/Formamidinium Lead Iodide Perovskite Solid Solutions from Temperature-Dependent Photoluminescence and Raman Spectroscopies," J. Phys. Chem. C **124**(6), 3448–3458 (2020).

[45] K. Opała, M. Saski, W. Marynowski, A. Borkenhagen, and J. Lewiński, "Exploring Polytype Formation in Double-Cation $DMA_xFA_{1-x}PbI_3$ Perovskites: Compositional Engineering and Phase Transitions," Chem. Mater. **37**(3), 897–911 (2025).

[46] C.J. Bartel, C. Sutton, B.R. Goldsmith, R. Ouyang, C.B. Musgrave, L.M. Ghiringhelli, and M. Scheffler, "New tolerance factor to predict the stability of perovskite oxides and halides," Sci. Adv. **5**(2), eaav0693 (2019).

[47] A. Kumar, A. Singh, and A.K. Ojha, "A new approach to predict the formation of 3D hybrid organic-inorganic perovskites," Int J of Quantum Chemistry **119**(22), e26012 (2019).

[48] H. Park, R. Mall, F.H. Alharbi, S. Sanvito, N. Tabet, H. Bensmail, and F. El-Mellouhi, "Learn-and-Match Molecular Cations for Perovskites," J. Phys. Chem. A **123**(33), 7323–7334 (2019).

SUPPLEMENTARY INFORMATION

Empirical equations to estimate tetragonal to cubic phase transitions in lead iodide perovskites

Fernando B. Minussi[1,2,3*], Rogerio M. Silva Jr.[4], Thays A. Abreu[5], Jose A. Eiras[3], Eudes B. Araujo[2]

[1] Department of Mechanical Engineering, Santa Catarina State University, Joinville, SC, Brazil

[2] Department of Physics and Chemistry, São Paulo State University, Ilha Solteira, SP, Brazil

[3] Department of Physics, Federal University of São Carlos, São Carlos, SP, Brazil

[4] Department of Electrical Engineering, São Paulo State University, Ilha Solteira, SP, Brazil

[5] Department of Mathematics, São Paulo State University, Ilha Solteira, SP, Brazil

Contents

Supplementary Note 1: Compositions and phase data of perovskites from literature _____ 1

Supplementary Note 2: Models' construction _____ 5

Supplementary Note 3: Visual comparison between MLA and SISSO results _____ 7

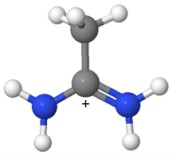
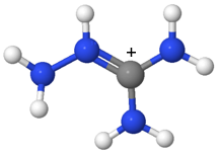
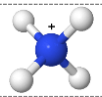
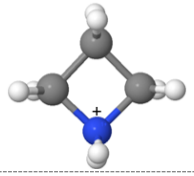
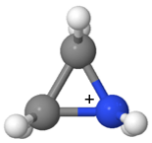
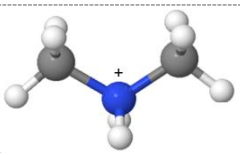
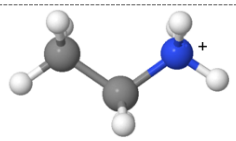
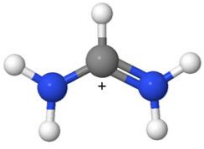
References _____ 8

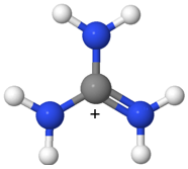
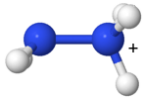
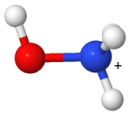
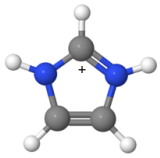
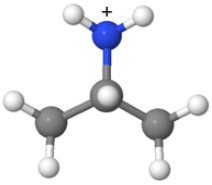
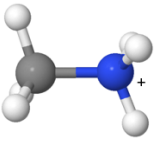
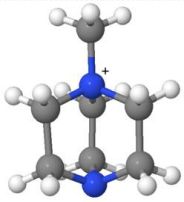
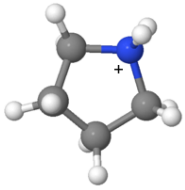
* fbminussi@gmail.com / f.minussi@unesp.br

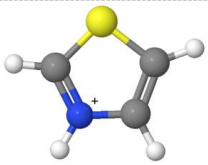
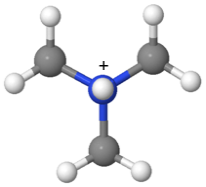
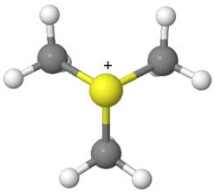
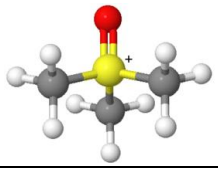
Supplementary Note 1: Compositions and phase data of perovskites from literature

The database available at https://tinyurl.com/3j9a22up contains phase data for 470 entries, some of which have the same compositions but come from different references. All compositions are of the $APbI_3$ type, containing single or mixed A-site with the chemical species given in Table S1. The characteristics r_A, μ_A, and NH_A were used to obtain the effective mixture descriptors, respectively r_{eff}, μ_{eff}, and n_{NH}, as given in Table 1 of main text.

Table S1 - Basic characteristics of all A-site cations contained in compounds of the data bank. Dipole moments estimated using MolCalc [1]. Color of spheres representing the elements in molecular cations: hydrogen (white), carbon (gray), nitrogen (blue), oxygen (red), and sulfur (yellow).

Cation (abbreviation)	Structure	r_A (pm)	μ_A (D)	NH_A	Ref.	Observations
Acetamidinium (AC^+)		277	2.02	4	[2]	
Aminoguanidinium (AGA^+)		351	2.22	7		Radius estimated with the procedure given by [3].
Ammonium (NH_4^+)		146	0	4	[3]	
Azetidinium (AZE^+)		250	3.79	2	[3]	
Aziridinium (AZI^+)		227	2.77	2	[4]	
Cesium (Cs^+)		188	isotropic	0	[5]	Ionic radius for CN = 12.
Dimethylammonium (DMA^+)		272	1.93	2	[3]	
Ethylammonium (EA^+)		274	4.97	2	[3]	
Formamidinium (FA^+)		253	0.21	4	[3]	

Guanidinium (GA^+)		278	0	6	[3]	
Hydrazinium (HY^+)		217	3.5	5	[3]	
Hydroxylammonium (HOA^+)		216	4.21	4	[3]	The O-H bond was included as a N-H bond.
Imidazolium (IM^+)		258	1.99	2	[3]	
Isopropylammonium (iPA^+)		282	5.57	3		For simplicity, we considered its radius the same as trMA.
Lithium (Li^+)		125	isotropic	0		Radius linearly extrapolated for CN = 12 from values for other CNs given by [5].
Methylammonium (MA^+)		217	2.7	3	[3]	
Methylhydrazinium (MHY^+)		264	5.13	4	[6]	
N-methyldabconium ($MDAB^+$)		375	5.55	0		Radius estimated with the procedure given by [3] with data from [7].
Potassium (K^+)		164	isotropic	0	[5]	Ionic radius for CN = 12.
Pyrrolidinium (Pyr^+)		272	4.77	2		Radius was considered the same of pyrrolinium [8].
Rubidinium (Rb^+)		172	isotropic	0	[5]	Ionic radius for CN = 12.

Thiazolium (Tz^+)		320	4.06	1	[8]	
Trimethylammonium ($trMA^+$)		282	1.12	0		Radius was considered as the mean between DMA^+ and tetramethylammonium (292 pm [3]).
Trimethylsulfonium (TMS^+)		244	1.07	0	[9]	
Trimethylsulfoxonium ($TMSO^+$)		289	5.77	0	[10]	

Data from 68 references were not included due to unavailable or insufficient phase information for the present work. Reasons for data not being useful include information collected under unspecified experimental conditions, very restricted measurement ranges, provided figures that lead to ambiguous results, results obtained by computational methods, and synthesis methods that are difficult to control. Some articles were also not included because they contained similar or identical data to others already extensively included in the database. Additionally, some systems may have been obtained under "unusual" conditions, such as in [11], leading to results that differ from the literature. Other systems, particularly $CsPbI_3$ and $FAPbI_3$ [12] and, more recently, $EAPbI_3$ [13,14], may not generate perovskites under usual synthesis conditions, but can form perovskites if other routes are employed. A recent example can be found at [15]. Furthermore, it is important to emphasize that films and nanocrystals may behave differently compared to large crystals. An example is $FA_xCs_{1-x}PbI_3$ nanocrystals, which appear to be total solid solutions [16], while in films and bulk crystals, phase separation is observed. Finally, despite earlier propositions [17], it seems to be mostly accepted that the sole or concomitant formation of the so-called $HPbI_3$ in halide perovskite systems is not feasible [18]. Hence, we opted to exclude its possibilities. Others, such as a morphotropic phase boundary in $MA_xFA_{1-x}PbI_3$ [19] were not considered either. Data from ten references were not included in the database because they contain data on materials in which the components of A-site compounds were not necessarily added to replace the A-site cations in the perovskite. Furthermore, the additives may have been used after the initial film formation. Therefore, it is not possible to infer whether the additional phases are due to segregated phases because they are systems above a

solubility limit, incomplete dissolution for kinetic reasons, etc. In other cases, additives containing cations or anions that may have been used (e.g. [20]) where it is not possible to guarantee that they have been eliminated from the final materials.

Supplementary Note 2: Models' construction

From the 470 entries, 125 are single-phased perovskites with existing tetragonal-to-cubic phase transition temperatures in the range of 200-350 K. These data were used for a multivariate linear regression analysis with the effective compositional descriptors as independent variables and the transition temperature as the dependent variable. The resulting equation in the form $T_{\mathrm{T}\to\mathrm{C}}\ (\mathrm{K}) = a + b(r_{\mathrm{eff}}) + c(\mu_{\mathrm{eff}}) + d(n_{\mathrm{NH}}) + e(\Delta S_{\mathrm{mix}})$ was used to calculate the $T_{\mathrm{T}\to\mathrm{C}}$ value for the compositions using the obtained a, b, c, d, and e coefficients. The experimental *versus* calculated values were plotted, and a linear fit with a slope equal to 1 was performed. As mentioned in the main text, eight points were excluded based on the criterion $|T_{\mathrm{calc}} - T_{\mathrm{exp}}|/T_{\mathrm{exp}} \geq 0.1$. The procedure was then redone to obtain a new, more accurate $T_{\mathrm{T}\to\mathrm{C}}$ polynomial.

To build the SISSO model, 129 data were used, including the same 125 plus four other compositions with tetragonal-to-cubic phase transition temperatures outside the 200 to 350 K range. Briefly, the SISSO approach is an iterative procedure that finds equations to describe a given set of dependent and independent variables through searching for the best R^2 model of pre-chosen operators applied to the initial descriptors. An operator applied to a descriptor makes a new descriptor and so on. When choosing the model complexity, the constructed code allows modifying two parameters:

nNonzeroCoefs = d; d = 1, 2, 3 or 4 - try models up to a maximum of 4D (4 terms).

nFeaturesPerSisIter = k; consider k new candidates per iteration.

The code only works if $d * k$ is lower than or equal to the number of descriptors generated. Usually, d = 3 (3D) or 4 (4D) and k =10 or 15 are used. For situations with a small number of descriptors, it is advisable to manipulate the values of the variables d and k so that their product is always less than or equal to the length of the Xout array. To perform SISSO regression in MATLAB using the Statistics and Machine Learning Toolbox, two files must be compiled (generateDescriptors.m and SissoRegressor.m), which are found in two independent scripts available at the following link: https://github.com/NREL/SISSORegressor_MATLAB/tree/main.

The function generateDescriptors(xCell, xVars) has the inputs xCell (a matrix of input feature cells, where each cell is a matrix of features with the same units) and xVars (variable names for the input features), and the class SissoRegressor(nNonzeroCoefs, nFeaturesPerSisIter) has the inputs nNonzeroCoefs (number of descriptors in the equation) and nFeaturesPerSisIter (number of iterations).

First, the database must be expanded with the combination of new descriptors (combined operations) that are called in the generateDescriptors script as operatorA1, operatorA2, and so go, according to the following list:

A1 - Square root

A2 - Square power

A3 - Cube power

A4 - Cube root

A5 - Fourth root

A6 - Fourth power

A7 - Fifth root

A8 - Fifth power

B1 - Inverse

C1 - Product of features from different groups

C2 - Absolute difference between features from the same group

C3 - Sum of features

Variables C2 and C2 were only enabled when there were different descriptors in the same unit of measurement. With the feature matrix expanded with the combination of operators, regression is performed using SISSO. Changes had to be made to the generateDescriptors.m file to include the C3 operator and update the syntax. The code used is given in: https://tinyurl.com/4euzpwpa. The entire set of models obtained is given in: https://tinyurl.com/ms3nukhs. For the present study, we opted to use the 2D relation with nFeaturesPerSisIter = 15 with operators A1, A2, A3, A4, A5, A6, B1, and C1, because it is quite simple and with a reasonably good RMSE, the best among the 2D models.

Supplementary Note 3: Visual comparison between MLA and SISSO results

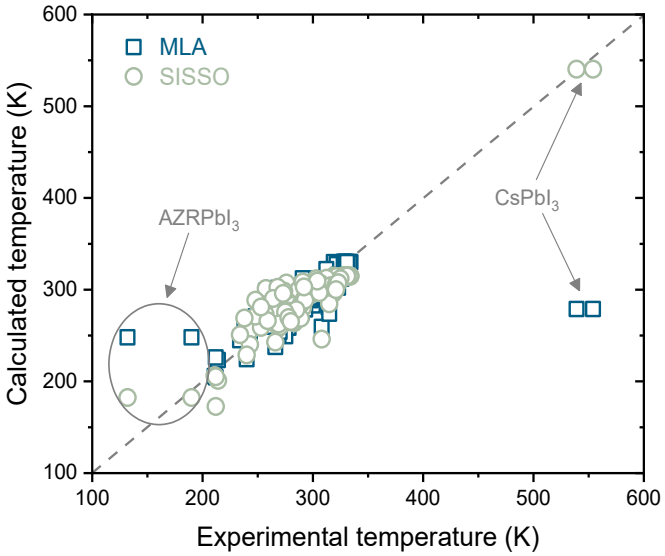


Figure S1 - Calculated temperatures for tetragonal to cubic phase transitions using the two models discussed in the main text. The highlighted data clearly show that MLA performs poorly for compositions with experimental transitions outside the 200-350 K range, but is generally better for data inside this range.

References

[1] J.H. Jensen, and J.C. Kromann, "The Molecule Calculator: A Web Application for Fast Quantum Mechanics-Based Estimation of Molecular Properties," J. Chem. Educ. **90**(8), 1093–1095 (2013).

[2] P. Singh, R. Mukherjee, and S. Avasthi, "Acetamidinium-Substituted Methylammonium Lead Iodide Perovskite Solar Cells with Higher Open-Circuit Voltage and Improved Intrinsic Stability," ACS Appl. Mater. Interfaces **12**(12), 13982–13987 (2020).

[3] G. Kieslich, S. Sun, and A.K. Cheetham, "Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog," Chem. Sci. 5(12), 4712–4715 (2014).

[4] C. Zheng, and O. Rubel, "Aziridinium Lead Iodide: A Stable, Low-Band-Gap Hybrid Halide Perovskite for Photovoltaics," J. Phys. Chem. Lett. **9**(4), 874–880 (2018).

[5] R.D. Shannon, "Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides," Acta Cryst A **32**(5), 751–767 (1976).

[6] M. Simenas, A. Gagor, J. Banys, and M. Maczka, "Phase Transitions and Dynamics in Mixed Three- and Low-Dimensional Lead Halide Perovskites," Chem. Rev. **124**(5), 2281–2326 (2024).

[7] B. Saparov, and D.B. Mitzi, "Organic–Inorganic Perovskites: Structural Versatility for Functional Materials Design," Chem. Rev. **116**(7), 4558–4596 (2016).

[8] G. Kieslich, S. Sun, and A.K. Cheetham, "An extended Tolerance Factor approach for organic–inorganic perovskites," Chem. Sci. **6**(6), 3430–3433 (2015).

[9] A. Kaltzoglou, C.C. Stoumpos, A.G. Kontos, G.K. Manolis, K. Papadopoulos, K.G. Papadokostaki, V. Psycharis, C.C. Tang, Y.-K. Jung, A. Walsh, M.G. Kanatzidis, and P. Falaras, "Trimethylsulfonium Lead Triiodide: An Air-Stable Hybrid Halide Perovskite," Inorg. Chem. **56**(11), 6302–6309 (2017).

[10] M. Parashar, R. Singh, K. Yoo, and J.-J. Lee, "Formation of 1-D/3-D Fused Perovskite for Efficient and Moisture Stable Solar Cells," ACS Appl. Energy Mater. **4**(3), 2751–2760 (2021).

[11] H. Park, A. Ali, R. Mall, H. Bensmail, S. Sanvito, and F. El-Mellouhi, "Data-driven enhancement of cubic phase stability in mixed-cation perovskites," Mach. Learn.: Sci. Technol. **2**(2), 025030 (2021).

[12] S. Hu, A.R. Tapa, X. Zhou, S. Pang, M. Lira-Cantu, and H. Xie, "Formation and stabilization of metastable halide perovskite phases for photovoltaics," Cell Reports Physical Science **5**(2), 101825 (2024).

[13] Y. Zhang, and N.-G. Park, "A thin film (<200 nm) perovskite solar cell with 18% efficiency," J. Mater. Chem. A **8**(34), 17420–17428 (2020).

[14] C.M. Guvenc, S. Toso, Y.P. Ivanov, G. Saleh, S. Balci, G. Divitini, and L. Manna, "Breaking the Boundaries of the Goldschmidt Tolerance Factor with Ethylammonium Lead Iodide Perovskite Nanocrystals," ACS Nano **19**(1), 1557–1565 (2025).

[15] S. Kralj, K. Artuk, A. Wieczorek, N. Orlov, Z. Eftekhari, R. Saive, E. Garnett, S. Siol, C.M. Wolff, and M. Morales-Masis, "Template-Assisted Growth of $Cs_x FA_{1-x} PbI_3$ with Pulsed Laser Deposition for Single Junction Perovskite Solar Cells," Advanced Energy Materials **15**(24), 2406033 (2025).

[16] J.A. Vigil, A. Hazarika, J.M. Luther, and M.F. Toney, "$FA_x Cs_{1-x} PbI_3$ Nanocrystals: Tuning Crystal Symmetry by A-Site Cation Composition," ACS Energy Lett. **5**(8), 2475–2482 (2020).

[17] F. Wang, H. Yu, H. Xu, and N. Zhao, "$HPbI_3$: A New Precursor Compound for Highly Efficient Solution-Processed Perovskite Solar Cells," Adv Funct Materials **25**(7), 1120–1126 (2015).

[18] W. Ke, I. Spanopoulos, C.C. Stoumpos, and M.G. Kanatzidis, "Myths and reality of HPbI3 in halide perovskite solar cells," Nat Commun **9**(1), 4785 (2018).

[19] T. Hainer, E. Fransson, S. Dutta, J. Wiktor, and P. Erhart, "A morphotropic phase boundary in MA1−xFAxPbI3: linking structure, dynamics, and electronic properties," Nat Commun **16**(1), 8775 (2025).

[20] N. Cheng, W. Li, M. Zhang, H. Wu, S. Sun, Z. Zhao, Z. Xiao, Z. Sun, W. Zi, and L. Fang, "Enhance the performance and stability of methylammonium lead iodide perovskite solar cells with guanidinium thiocyanate additive," Current Applied Physics **19**(1), 25–30 (2019).

Empirical equations to estimate band gap energies in halide perovskites

Fernando B. Minussi[1,2,3*], Rogerio M. Silva Jr.[4], Thays A. Abreu[5], Jose A. Eiras[3], Eudes B. Araujo[2]

[1] Department of Mechanical Engineering, Santa Catarina State University, Joinville, SC, Brazil

[2] Department of Physics and Chemistry, São Paulo State University, Ilha Solteira, SP, Brazil

[3] Department of Physics, Federal University of São Carlos, São Carlos, SP, Brazil

[4] Department of Electrical Engineering, São Paulo State University, Ilha Solteira, SP, Brazil

[5] Department of Mathematics, São Paulo State University, Ilha Solteira, SP, Brazil

Para este trabalho, uma metodologia análoga à do anteriormente apresentado foi aplicada no banco de dados organizado em https://tinyurl.com/3w2dv2y9. O arquivo contém energia de *band gap* de 464 entradas, algumas das quais possuem as mesmas composições, mas que vêm de referências diferentes. As composições químicas, envolvendo os componentes da Tabela 1 abaixo, foram convertidas em seus descritores efetivos ("*averaged descriptors*") e, então, nos parâmetros t (fator de tolerância de Goldschmidt), μ (distorção octaedral) e IC (caráter iônico), conforme mostrado na Tabela 2. Utilizando-se regressão linear multivariada (MLR) e a metodologia SISSO foram obtidos dois modelos empíricos, também mostrados na Tabela 2. A relação entre dados experimentais e calculados com esses modelos são dados na Figura 1. Apesar da simplicidade dos modelos, o grande trunfo desse trabalho é fornecer modelos capazes de estimar as energias de *band gap* de composições ABX_3 quaisquer de perovskitas de haleto a partir de informações facilmente determinadas e com boa acurácia, sendo os erros da ordem de apenas 0,18 eV. Um trabalho dessa forma inexiste na literatura atual.

* fbminussi@gmail.com / f.minussi@unesp.br

Tabela 1 - Chemical information of the A-, B-, and X-site components of the compositions given in the databank.

A-SITE CATIONS			
Name	**Abbreviation**	**Radius (pm)**	**Reference**
cesium	Cs^+	188	1
methylammonium	MA^+	217	2
formamidinium	FA^+	253	2
ethylammonium	EA^+	274	2
guanidinium	GA^+	278	2
acetamidinium	AC^+	277	3
dimethylammonium	DMA^+	272	2
aziridinium	AZR^+	227	4
methylhydrazinium	MHy^+	264	5
methylphosphonium	MP^+	220	#
trimethylammonium	TMA^+	282	6
rubidium	Rb^+	172	1
azetidinium	AZE^+	250	2
ethylenediammonium	en^{2+}	333	7
B-SITE CATIONS			
Abbreviation	**Electronegativity***	**Radius (pm)**	**Reference**
Pb^{2+}	1.87	119	1
Sn^{2+}	1.96	115	
Ge^{2+}	2.01	73	
Cu^{2+}	1.9	73	
Ca^{2+}	1.00	100	
Cd^{2+}	1.69	95	
X-SITE ANION			
Abbreviation	**Electronegativity***	**Radius (pm)**	**Reference**
I^-	2.66	220	1
Br^-	2.96	196	
Cl^-	3.16	180	

* Tabulated elemental Pauling electronegativity values.
# Estimated with the procedure given in [2].

Tabela 2 - Summary of the conversion equations from chemical descriptors, parameter relations from the averaged descriptors and derived empirical equations from literature data using multivariate linear analysis (MLR) and sure independence screening and sparsifying operator (SISSO) approaches. a_i, b_i, and x_i and r_A, r_B, and r_X are the molar fraction and ionic radius of each A-, B-, and X-site components, respectively. χ_B and χ_X are the electronegativity each B- and X-site components, respectively. RMS is the root mean squared error.

Averaged descriptors	$R_A = \sum_{i=1}^{\infty} a_i(r_{A,i})$	$R_B = \sum_{i=1}^{\infty} b_i(r_{B,i})$	$R_X = \sum_{i=1}^{\infty} x_i(r_{X,i})$	$EN_B = \sum_{i=1}^{\infty} b_i(\chi_{B,i})$	$EN_X = \sum_{i=1}^{\infty} x_i(\chi_{X,i})$
Parameters	$t = \frac{R_A + R_X}{\sqrt{2}(R_B + R_X)}$	$\mu = \frac{R_B}{R_X}$	$IC = 1 - \exp\left[-\frac{(EN_B - EN_X)^2}{4}\right]$		
MLR	$E_g = 1.216 + 0.785(t) - 3.508(\mu) + 10.040(IC)$	RMS = 0.180 eV			
SISSO	$E_g = 0.553 + 2.967\left(\frac{IC}{\mu}\right) + 9.648\,[(t)(IC)]^2$	RMS = 0.177 eV			

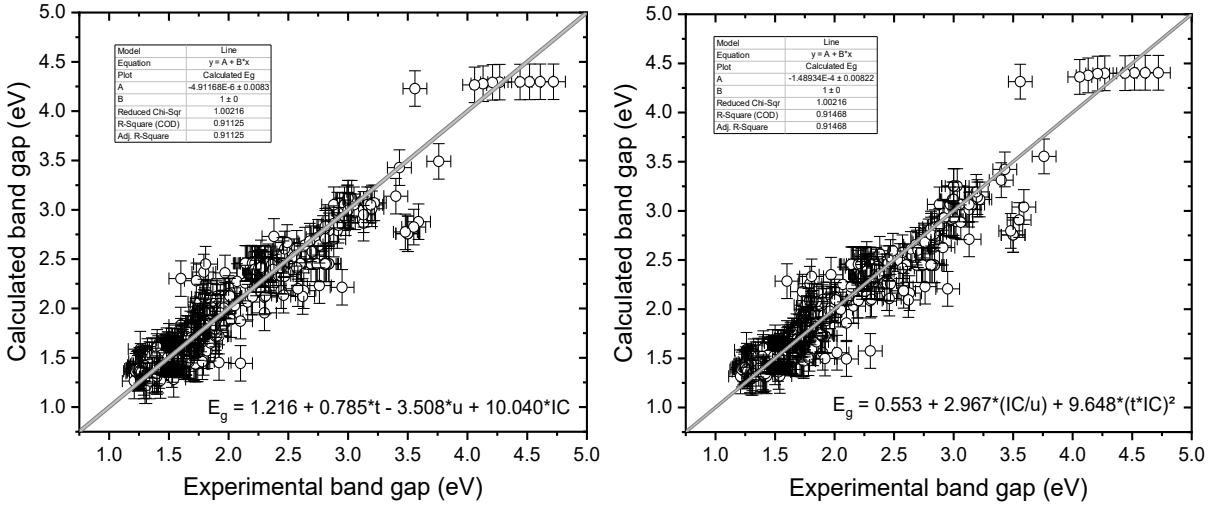


Figura 1 - Correlação entre valores experimentais e calculados usando os métodos MLR (esquerda) e SISSO (direita).

Referências

[1] R.D. Shannon, "Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides," Acta Cryst A **32**(5), 751–767 (1976).

[2] G. Kieslich, S. Sun, and A.K. Cheetham, "Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog," Chem. Sci. 5(12), 4712–4715 (2014).

[3] P. Singh, R. Mukherjee, and S. Avasthi, "Acetamidinium-Substituted Methylammonium Lead Iodide Perovskite Solar Cells with Higher Open-Circuit Voltage and Improved Intrinsic Stability," ACS Appl. Mater. Interfaces **12**(12), 13982–13987 (2020).

[4] C. Zheng, and O. Rubel, "Aziridinium Lead Iodide: A Stable, Low-Band-Gap Hybrid Halide Perovskite for Photovoltaics," J. Phys. Chem. Lett. **9**(4), 874–880 (2018).

[5] M. Simenas, A. Gagor, J. Banys, and M. Maczka, "Phase Transitions and Dynamics in Mixed Three- and Low-Dimensional Lead Halide Perovskites," Chem. Rev. **124**(5), 2281–2326 (2024).

[6] Empirical equations to estimate tetragonal to cubic transitions in lead iodide perovskites (to be submitted)

[7] I. Spanopoulos, W. Ke, C.C. Stoumpos, E.C. Schueller, O.Y. Kontsevoi, R. Seshadri, and M.G. Kanatzidis, "Unraveling the Chemical Nature of the 3D 'Hollow' Hybrid Halide Perovskites," J. Am. Chem. Soc. **140**(17), 5728–5742 (2018).