

# The acidic intrinsically disordered region of the inflammatory mediator HMGB1 mediates fuzzy interactions with CXCL12

Received: 17 May 2023

Accepted: 26 January 2024

Published online: 08 February 2024

 Check for updates

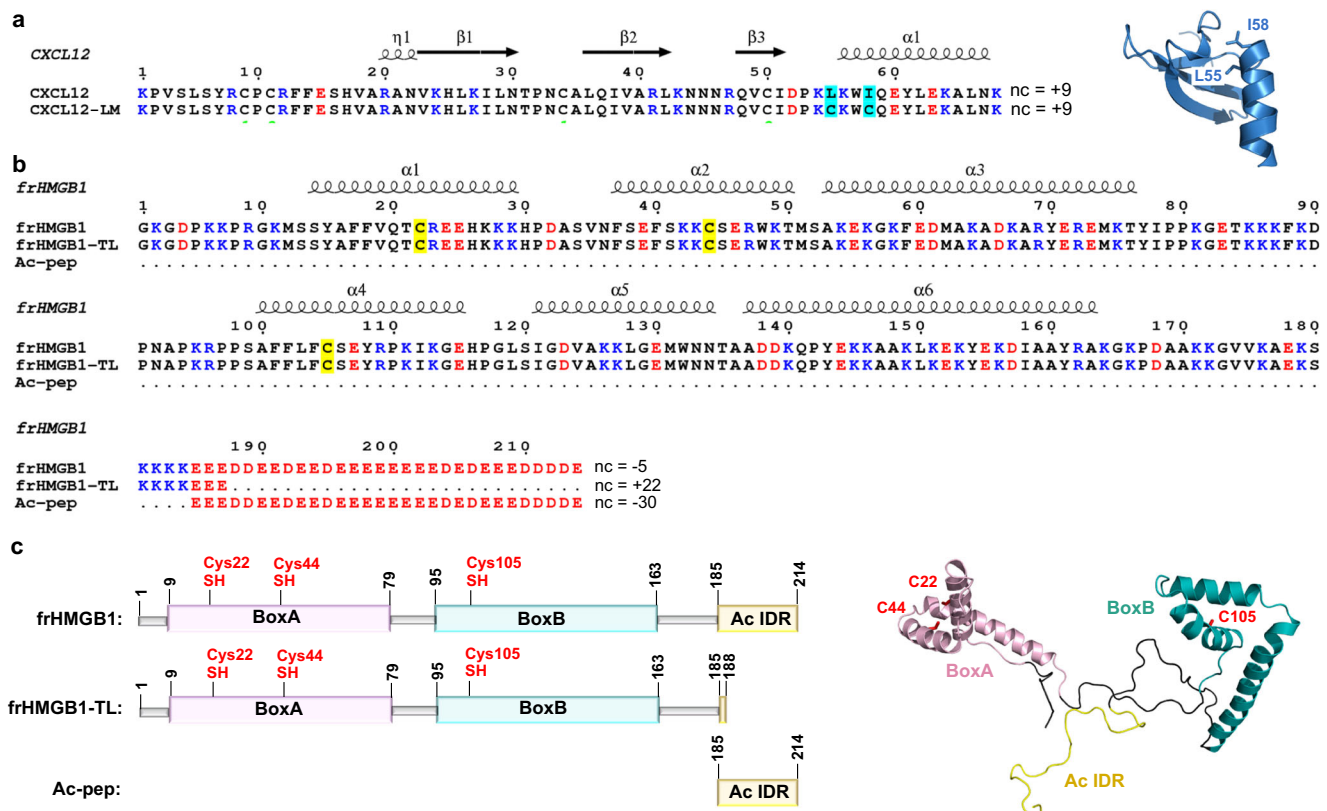
Malisa Vittoria Mantonico <sup>1,2</sup>, Federica De Leo<sup>1,9</sup>, Giacomo Quilici<sup>1</sup>, Liam Sean Colley <sup>3,4</sup>, Francesco De Marchis<sup>2,5</sup>, Massimo Crippa <sup>5</sup>, Rosanna Mezzapelle <sup>2,5</sup>, Tim Schulte<sup>6</sup>, Chiara Zucchelli<sup>1</sup>, Chiara Pastorello <sup>1</sup>, Camilla Carmeno <sup>1</sup>, Francesca Caprioglio <sup>2,5</sup>, Stefano Ricagno<sup>6,7</sup>, Gabriele Giachin <sup>8</sup>, Michela Ghitti <sup>1</sup>✉, Marco Emilio Bianchi <sup>2,5</sup> & Giovanna Musco <sup>1</sup>✉

Chemokine heterodimers activate or dampen their cognate receptors during inflammation. The CXCL12 chemokine forms with the fully reduced (fr) alarmin HMGB1 a physiologically relevant heterocomplex (frHMGB1•CXCL12) that synergically promotes the inflammatory response elicited by the G-protein coupled receptor CXCR4. The molecular details of complex formation were still elusive. Here we show by an integrated structural approach that frHMGB1•CXCL12 is a fuzzy heterocomplex. Unlike previous assumptions, frHMGB1 and CXCL12 form a dynamic equimolar assembly, with structured and unstructured frHMGB1 regions recognizing the CXCL12 dimerization surface. We uncover an unexpected role of the acidic intrinsically disordered region (IDR) of HMGB1 in heterocomplex formation and its binding to CXCR4 on the cell surface. Our work shows that the interaction of frHMGB1 with CXCL12 diverges from the classical rigid heterophilic chemokines dimerization. Simultaneous interference with multiple interactions within frHMGB1•CXCL12 might offer pharmacological strategies against inflammatory conditions.

Chemokines constitute a large family of signaling proteins that interact with cell surface chemokine G protein-coupled receptors (GPCRs). Through an intricate network of cross-talks with their receptors they regulate leukocyte activation and trafficking in physiological and pathological conditions<sup>1,2</sup>. Chemokines are structurally characterized by the presence of two conserved disulfide bridges, a flexible N-terminus, a three-stranded  $\beta$ -sheet packing on a C-terminal  $\alpha$ -helix and

3 loops connecting the elements of secondary structure. They usually exist in a monomer-multimer equilibrium. In particular, members of the CXC subfamily, like CXCL12 (C-X-C motif ligand 12), can dimerize via intermonomer contacts between the first  $\beta$ -strand and the  $\alpha$ -helix of each monomer (Fig. 1a, Supplementary Fig. 1a). A shift towards one or the other form can either activate or dampen their cognate receptors, thus adding a further layer of complexity to the functional tuning

<sup>1</sup>Biomolecular NMR Laboratory, Division of Genetics and Cell Biology, IRCCS Ospedale San Raffaele, Milan, Italy. <sup>2</sup>School of Medicine, Università Vita e Salute-San Raffaele, Milan, Italy. <sup>3</sup>HMGBiotech S.r.l., 20133 Milan, Italy. <sup>4</sup>School of Medicine and Surgery, Università Milano-Bicocca, 20126 Milan, Italy. <sup>5</sup>Chromatin Dynamics Unit, Division of Genetics and Cell Biology, IRCCS Ospedale San Raffaele, Milan, Italy. <sup>6</sup>Institute of Molecular and Translational Cardiology, IRCCS Policlinico San Donato, Milan, Italy. <sup>7</sup>Department of Biosciences, Università degli Studi di Milano, Milan, Italy. <sup>8</sup>Department of Chemical Sciences (DiSC), University of Padua, 35131 Padova, Italy. <sup>9</sup>Present address: Experimental Therapeutics Program, IFOM ETS - The AIRC Institute of Molecular Oncology and AIRC, Fondazione AIRC per la Ricerca sul Cancro ETS, Milan, Italy. ✉e-mail: [ghitti.michela@hsr.it](mailto:ghitti.michela@hsr.it); [musco.giovanna@hsr.it](mailto:musco.giovanna@hsr.it)



**Fig. 1 | CXCL12 and frHMGB1 constructs.** Amino acid sequence of: (a) CXCL12 and locked monomer CXCL12 mutant L55C/I58C (CXCL12-LM) and (b) frHMGB1, frHMGB1-TL and Ac-pep. Basic and acidic residues are shown in blue and red, respectively, nc indicates the net charge. L55C and I58C are colored in cyan on CXCL12 sequences. C22, C44 and C105 are colored in yellow on HMGB1 sequences. On the top of the alignments the elements of secondary structures are indicated. In (a) on the right is reported the cartoon representation of CXCL12 (pdb code: 2KEE),

with L55 and I58 explicitly shown in sticks. c Schematic diagram of the fully reduced form of HMGB1 (frHMGB1), tail-less frHMGB1 (frHMGB1-TL), the acidic peptide (Ac-pep) corresponding to the HMGB1 acidic intrinsically disordered region (Ac IDR). BoxA (pink), BoxB (cyan) and Ac IDR (yellow) are represented with boxes and colored on frHMGB1 structure (AF-P63159). The side chains of fully reduced cysteines are represented in red sticks.

of chemokine/receptor axis and of their downstream pathways<sup>3-5</sup>. Another sophisticated mechanism of chemokine fine-regulation is their ability to heterodimerize with other chemokines, resulting in synergic activation and dimerization of their cognate receptors<sup>6-8</sup>. This mechanism is particularly relevant in inflammatory conditions, where chemokine heterocomplexes activate chemokine receptors in the presence of low concentrations of chemokine-selective agonists that otherwise, without the synergy-inducing partner, would be inactive<sup>8,9</sup>. Examples include the interaction of: (i) CCL21 or CCL19 with CXCL13 to enhance leukocyte migration and activities through binding and activation of CCR7 at lower agonist concentrations<sup>10</sup>, (ii) CCL5 with CXCL4 to recruit monocytes and neutrophils<sup>11</sup>, and (iii) CXCL12 and CXCL9 with CXCR4 to attract lymphoma cells<sup>12</sup>. Such a chemokine interactome further enlarges the possibility of heteromeric interactions among chemokines, expanding the fine-regulation of signaling possibilities<sup>11,13</sup>.

Intriguingly, chemokines can also form heterophilic interactions with some inflammatory mediators that are not structurally homologous to the classical CC-, CXC-, CX3C-, or XC-chemokines<sup>4</sup>. In this sense CXCL12 represents a paradigmatic example, as it is also able to bind to other proteins, such as galectins<sup>14</sup> and the alarmin High Mobility Group Box 1 (HMGB1)<sup>6,15</sup>. The interaction with Galectin 3 is immunoregulatory and attenuates CXCL12-stimulated signaling via CXCR4<sup>14</sup>. Conversely, binding of CXCL12 to HMGB1 synergistically enhances the CXCR4-dependent chemotactic response of monocytes and is involved in tissue regeneration, cell proliferation and tumour progression<sup>6,15,16</sup>. The 25 kDa HMGB1 protein comprises two L-shaped HMG tandem boxes (~80 aa each connected by a flexible linker),

referred to as BoxA and BoxB, and an acidic C-terminal intrinsically disordered region (IDR, 30 aa) (Fig. 1b, c). HMGB1 is a Damage Associated Molecular Pattern (DAMP), which once released in the extracellular space alerts the host to stress, unscheduled cell death or microbial invasion, thus triggering inflammation and immune responses<sup>17,18</sup>. HMGB1 is passively released by dead non-apoptotic cells and is actively released by severely stressed cells and by immune cells such as macrophages, natural killer cells, neutrophils and mature dendritic cells (reviewed in<sup>19</sup>). HMGB1 contains three cysteines (C22, C44, and C105) (Fig. 1b, c), whose redox states determine how it functions as a pro-inflammatory mediator<sup>20</sup>. On one hand, the disulfide form (dsHMGB1), with C22 and C44 forming a disulfide bridge on BoxA, binds to the Toll-like receptor 4/MD2 complex, herewith promoting inflammatory responses and cytokine activation<sup>21</sup>. On the other hand, the fully reduced form (frHMGB1) plays a pivotal role in promoting the recruitment of inflammatory cells to injured tissues via heterocomplex formation with CXCL12 (frHMGB1-CXCL12) and activation of CXCR4. This heterocomplex induces specific CXCR4 homodimer rearrangements, promotes CXCR4-mediated signaling, resulting in increased ERK activation and calcium rise induction<sup>15</sup> and maintains CXCR4 on the plasma membrane in a  $\beta$ -arrestin 2 dependent manner<sup>22</sup>. HMGB1, CXCL12 and CXCR4 are highly expressed in inflammation related cancers, like malignant mesothelioma<sup>23</sup>, where they contribute to disease initiation and progression<sup>16</sup>. Pharmacological targeting of the frHMGB1-CXCL12-CXCR4 axis is thus emerging as an appealing opportunity against inflammation related diseases<sup>24</sup>.

While Galectin 3 is structurally reminiscent of chemokines and interacts with CXCL12 exploiting in part the CXC-type dimerization

surface, composed by the beta-strand  $\beta 1$  and the alpha-helix  $\alpha 1^4$  (Fig. 1a, Supplementary Fig. 1a), the molecular details dictating frHMGB1•CXCL12 intermolecular interactions are in part elusive, possibly because of the intrinsic dynamics of the system components. On one side CXCL12, as a classical chemokine, exists in a physiologically relevant monomer-dimer equilibrium<sup>25</sup>. On the other side, HMGB1 is distinguished by conformational heterogeneity and variable contact patterns: it oscillates between a collapsed and open form, via intramolecular electrostatic interaction between the acidic IDR and the basic HMG boxes<sup>26–29</sup>. Whether one or two molecules of CXCL12 bind to HMGB1 and whether the acidic IDR plays a role in heterocomplex formation and function are still open questions<sup>5,30</sup>.

Here, to address these issues, we adopt a dissecting and mutagenesis strategy coupled to an integrative structural approach (NMR, AUC, ITC, MST, SAXS). Our findings uncover the formation of an equimolar complex between frHMGB1 and CXCL12, with the acidic IDR assuming a previously overlooked, yet crucial role in complex formation. The resulting complex is best described by a heterogeneous ensemble rather than by a single defined structure, with frHMGB1 retaining its intrinsic dynamics during its interaction with CXCL12. In particular, NMR titrations and <sup>15</sup>N relaxation experiments indicate that the acidic IDR contains multiple sites capable of interacting with CXCL12 and maintains its conformational flexibility when bound to CXCL12. Moreover, CXCL12 employs the same interaction surface to establish multivalent interactions with both BoxA and the acidic IDR. Importantly, we provide evidence that in a cellular context the acidic IDR facilitates the binding of the frHMGB1•CXCL12 heterocomplex to CXCR4 on the extracellular surface. All these features align well with the concept of fuzzy interactions, where at least one component retains its dynamic nature and has a discernible impact on functional outcomes<sup>31,32</sup>.

## Results

### HMGB1 acidic IDR takes part in frHMGB1•CXCL12 formation

The molecular details and the actual role of the single HMGB1 domains in the formation of the heterocomplex with CXCL12 are still elusive. Existing models, based on NMR titrations between the single HMG boxes and CXCL12, are based on the assumption that only the structured domains are the main actors in complex formation<sup>30</sup>. Whether the acidic IDR of frHMGB1 plays a role in complex formation has never been explored. We have therefore adopted a dissecting approach and performed comparative nuclear magnetic resonance (NMR) experiments titrating <sup>15</sup>N CXCL12 with a synthetic peptide corresponding to the acidic IDR (Ac-pep), a tail-less HMGB1 construct, composed of the fully reduced HMG tandem domain devoid of the acidic IDR (frHMGB1-TL), and a fully reduced full-length HMGB1 (frHMGB1) (Fig. 1b, c).

Notably, addition of Ac-pep to <sup>15</sup>N CXCL12 induced a dramatic change of the corresponding <sup>1</sup>H-<sup>15</sup>N Heteronuclear Single Quantum Coherence (HSQC) spectrum, with substantial chemical shift perturbations (CSPs, in the 0.2–0.6 ppm range) and overall peak intensity reduction (Fig. 2a–c). The interaction occurred in the intermediate exchange on the NMR chemical shift time scale: peaks were severely broadened or disappeared beyond detection upon addition of sub-stoichiometric amounts of Ac-pep, but reappeared at 1:1 stoichiometry (Supplementary Fig. 2a). The dissociation constant was in the sub-micromolar range ( $K_d = 0.2 \pm 0.1 \mu\text{M}$ ), as assessed by line-shape analysis using the software TITAN 1.6<sup>33</sup> (Supplementary Fig. 2b). CXCL12 residues with the highest CSPs were R12, H17, V18, R20, V23, K24, H25, A40, R41, N45, W57, L66, N67 (Fig. 2b), with V23, K24, H25 (on  $\beta 1$ ) and L66, N67 (on  $\alpha 1$ ) all located on the known CXCL12 homodimerization interface (Fig. 2d, Supplementary Fig. 1a).

In contrast to Ac-pep, addition of frHMGB1-TL to <sup>15</sup>N CXCL12 (Fig. 2e–g) induced very small CSPs (in the 0.02–0.06 ppm range). As typically observed in NMR studies of chemokine heterocomplex formation, the interaction occurred again in the intermediate exchange

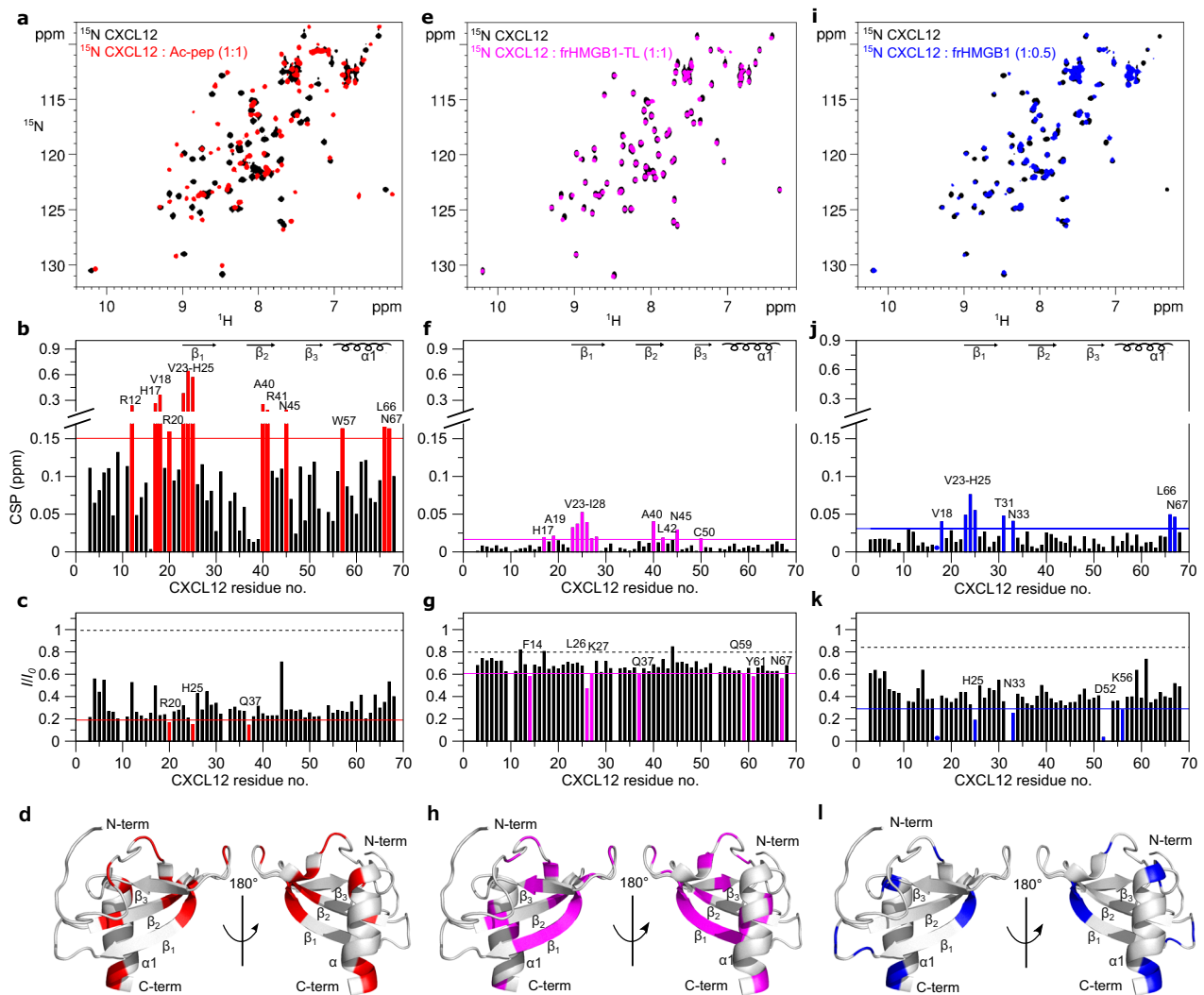
regime on the NMR chemical shift timescale<sup>14</sup>, with peak-intensity reduction upon binding. Residues mostly affected by the interaction with frHMGB1-TL partially coincided with the ones affected by Ac-pep (H17, V23-H25, A40, N45, N67), suggesting that both the acidic IDR and the HMG tandem domain in part share the same interaction surface (Fig. 2h). Similarly, titration of full-length frHMGB1 induced significant CSPs (in the 0.05–0.1 ppm range) on residues located on the  $\beta 1$  strand (V23-H25) and the  $\alpha$ -helix (L66, N67) (Fig. 2i, j). Herein, we observed pronounced line broadening effects already at sub-stoichiometric concentrations (1:0.5) (Fig. 2i) that hampered analysis at equimolar ratio, suggesting different binding dynamics and affinities for the full-length protein and its different constructs. Of note, comparison of the profiles of the CSPs of <sup>15</sup>N CXCL12 upon addition of either Ac-pep, frHMGB1-TL or frHMGB1 were similar, with residues located on the CXCL12 dimerization surface showing the highest CSPs (Fig. 2i). Importantly, the chemical shifts of these residues and their perturbations strongly depend on the CXCL12 monomer-dimer equilibrium<sup>5</sup>. Thus, to distinguish CSPs due to direct interactions with frHMGB1 and constructs thereof from those related to potential changes in the CXCL12 oligomerization state, we repeated the titrations with a CXCL12 mutant locked in a monomeric state (CXCL12-LM)<sup>34</sup>. This mutant, which restricts the  $\alpha 1$  region to a specific orientation relative to the  $\beta$ -sheet through a disulfide bond (between C55 and C58) (Fig. 1a, Supplementary Fig. 1b), was previously engineered with the aim to block chemokine oligomerization and to distinguish CSPs resulting from ligand induced dimerization from those stemming from direct interactions<sup>34,35</sup>. Upon the addition of Ac-pep, frHMGB1-TL, and frHMGB1, we observed that the spectral perturbations in CXCL12-LM were similar to those seen in the wild-type CXCL12 in terms of line broadening and CSP profiles. However, the CSPs magnitude of <sup>15</sup>N CXCL12-LM upon addition of frHMGB1-TL and frHMGB1 was larger than that observed in the wild-type protein, possibly due to different internal dynamics and/or binding kinetics in the wild-type and mutant protein. Nonetheless, the resemblance in CSP profiles strongly suggests that the N-terminal portion of the  $\beta 1$ -strand and the  $\alpha 1$ -helix of CXCL12 indeed constitute an authentic interaction surface (Supplementary Fig. 3).

Taken together, comparison of <sup>15</sup>N CXCL12 NMR titrations with frHMGB1 fragments revealed that the acidic IDR of HMGB1 is directly involved in frHMGB1•CXCL12 heterocomplex formation and that the CXCL12 dimerization surface works as hub for multivalent interactions with both the acidic IDR and the HMG tandem domain.

### The frHMGB1•CXCL12 heterocomplex forms via fuzzy interactions

We next hypothesized that the inter-molecular interactions between the acidic IDR and CXCL12 might perturb frHMGB1 intra-molecular interactions and conformational equilibria. Indeed, in reversed titrations, frHMGB1 amide resonances of spy residues reported to interact with the acidic IDR (e.g. W48, T76, I78 on BoxA, and A93, I158 on BoxB)<sup>28</sup>, moved towards their NMR frequency in the tailless construct upon addition of CXCL12. These displacements suggest a weakening of the intramolecular interactions between the acidic IDR and the frHMG boxes (Fig. 3a). The shifts towards frHMGB1-TL resonances were relatively small, presumably because CXCL12 only partially competes with frHMGB1 intra-molecular interactions. Addition of an equimolar amount of Ac-pep to <sup>15</sup>N frHMGB1 in complex with CXCL12 was sufficient to sequester CXCL12 and to disrupt the heterocomplex, as indicated by the reappearance of <sup>15</sup>N frHMGB1 resonances, confirming the important contribution of the acidic IDR to heterocomplex formation (Fig. 3b).

Overall, CSPs and intensity variations in <sup>15</sup>N frHMGB1/CXCL12 NMR titrations are due to both intra- and inter-molecular interactions, hence mapping of spectral perturbations on the frHMGB1 structure reflects both phenomena, which are difficult to separate (Fig. 3c–e).



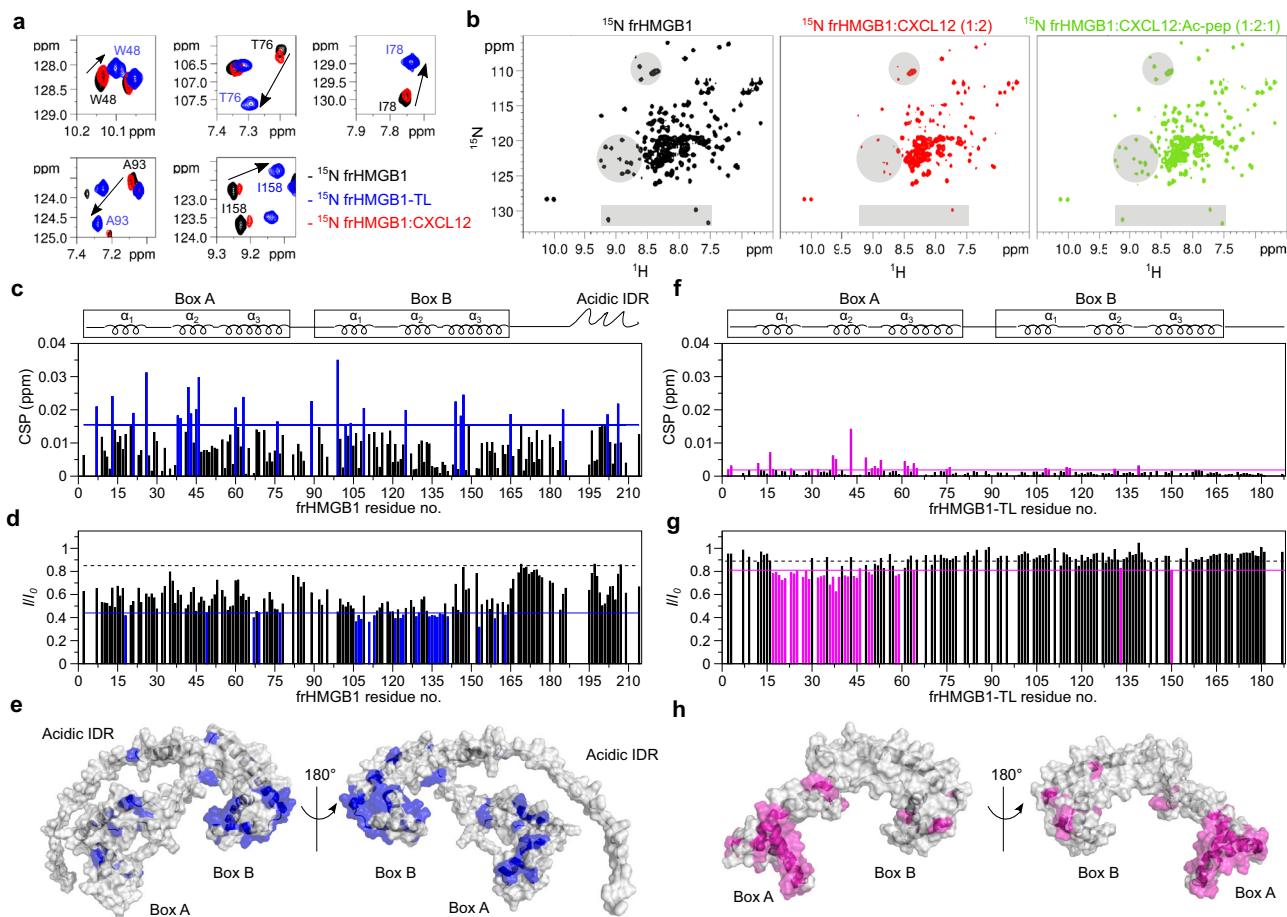
**Fig. 2 | Ac-pep, frHMGB1-TL and frHMGB1 interact with CXCL12 dimerization surface.** **a** Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$  CXCL12 (0.1 mM) without (black) and with (red) Ac-pep (1:1). Bar graph showing **(b)** residue-specific chemical shift perturbation (CSPs) and **(c)** peak intensities ratios ( $I/I_0$ ) of  $^{15}\text{N}$ -labeled CXCL12 (0.1 mM) upon addition of Ac-pep (1:1). Residues with  $\text{CSP} > \text{avg} + \sigma_0$  (corrected standard deviation, red line) and with  $I/I_0 < \text{avg} - \text{SD}$  (standard deviation, red line) are labeled and **(d)** shown in red on CXCL12 (gray cartoon, pdb code: 2KEE). **e** Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$  CXCL12 (0.1 mM) without (black) and with (magenta) frHMGB1-TL (1:1). Bar graph showing **(f)** residue-specific CSPs and **(g)**  $I/I_0$  of  $^{15}\text{N}$ -labeled CXCL12 (0.1 mM) upon addition of frHMGB1-TL (1:1) of  $^{15}\text{N}$ -labeled CXCL12 (0.1 mM). Residues with  $\text{CSP} > \text{avg} + \sigma_0$  (magenta line) with  $I/I_0$

$< \text{avg} - \text{SD}$  (magenta line) are labeled and **(h)** shown in magenta on CXCL12. **i** Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$  CXCL12 (0.1 mM) without (black) and with (blue) frHMGB1 (1:0.5). Bar graph showing **(j)** residue-specific CSPs and **(k)**  $I/I_0$  of  $^{15}\text{N}$ -labeled CXCL12 (0.1 mM) upon addition of frHMGB1 (1:0.5). Residues with  $\text{CSP} > \text{avg} + \sigma_0$  (blue line) and  $I/I_0 < \text{avg} - \text{SD}$  (blue line) are labeled and **(l)** shown in blue on CXCL12. In the bar graphs  $\alpha$ -helices and  $\beta$ -strands are schematically represented on the top, missing residues are prolines, dots indicate residues disappearing upon binding, the dashed black line indicates the expected peak intensity decrease due to the titration dilution effect. Source data are provided as a Source Data file.

Thus, to remove the confounding effect of the acidic IDR, we performed NMR titrations of  $^{15}\text{N}$  frHMGB1-TL with CXCL12. Notably, the CSPs and peak intensity reduction profiles were different and substantially smaller than the ones observed in the full-length protein. Nonetheless, the removal of the acidic IDR brought out the presence of an interaction surface formed by the first two helices of BoxA, with no major involvement of BoxB (Fig. 3f–h). The relevant role of BoxA in the interaction was further emphasized by the fact that addition of CXCL12 to the oxidized tailless construct ( $^{15}\text{N}$  dsHMGB1-TL), produced markedly reduced spectral changes (Supplementary Fig. 4a) compared to frHMGB1-TL (Fig. 3f, g). A similar result was obtained in the reverse titration (Supplementary Fig. 4b). This confirms that BoxA and its oxidation state are indeed crucial for the interaction with CXCL12 (ref. 20). Notably, and similar to frHMGB1, in NMR titrations of  $^{15}\text{N}$  dsHMGB1 and CXCL12 (Supplementary Fig. 4c) and reverse (Supplementary Fig. 4d), the major spectral perturbations were mainly due to

the intra- and inter-molecular interactions of the acidic IDR, whose binding to CXCL12 is independent from the redox state of BoxA.

The dynamic nature of the acidic IDR interaction with CXCL12 was further confirmed by NMR titrations of recombinant  $^{15}\text{N}$  acidic IDR (Ac-pep<sub>rec</sub>) with unlabelled CXCL12. As is typical for fuzzy interactions involving IDRs or intrinsically disordered proteins (IDPs)<sup>36,37</sup>, the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$  Ac-pep<sub>rec</sub>, in both the absence and presence of equimolar CXCL12, displayed reduced peak dispersion and high signal overlap, with minimal chemical shift perturbations (Fig. 4a). Similarly, H $\alpha$ -C $\alpha$  and H $\beta$ -C $\beta$  peak clusters of aspartic and glutamic residues in the  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectra did not show significant chemical shift changes upon CXCL12 binding, indicating the absence of persistent secondary structure formation (Fig. 4b). The pronounced overlap in the NMR spectra and the repetitive amino acid sequence prevented specific residue assignments. As a result, peaks were assigned arbitrary numbers, and the interpretation of their intensities should primarily be



**Fig. 3 | The frHMGB1-CXCL12 heterocomplex forms via fuzzy interactions.**

**a** Superposition of selected regions of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 0.1 mM  $^{15}\text{N}$  frHMGB1 (corresponding to spy residues W48, T76, I78, A93, I158) without (black) and with 0.2 mM CXCL12 (red) and 0.1 mM  $^{15}\text{N}$  frHMGB1-TL (blue). CXCL12 partially competes with intramolecular frHMGB1 interactions and specific amide resonances move (arrow) towards the chemical shift of the corresponding amide in the tailless construct. **b**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of frHMGB1 (0.1 mM) without (black) and with 0.2 mM CXCL12 (red), and with subsequent addition of 0.1 mM Ac-pep (green). Grey shadowed regions highlight resonances disappearing and reappearing upon addition of CXCL12 and Ac-pep, respectively. Bar graphs showing (c) residue-specific chemical shift perturbation (CSP) and (d) peak intensity ratios ( $I/I_0$ ) of  $^{15}\text{N}$ -labeled frHMGB1 (0.1 mM) upon addition of CXCL12 (1:1). Residues with

CSP > avg +  $\sigma_0$  (corrected standard deviation, blue line) and  $I/I_0$  < avg - SD (standard deviation, blue line) are colored (blue) and (e) mapped on frHMGB1 (grey surface, AphaFold2 model AF-P63159). Bar graph showing (f) residue-specific CSPs and (g)  $I/I_0$  of  $^{15}\text{N}$ -labeled frHMGB1-TL (0.1 mM) upon addition of CXCL12 (1:1). Residues with CSP > avg +  $\sigma_0$  (magenta line) and  $I/I_0$  < avg - SD (magenta line) are colored in magenta and (h) mapped on frHMGB1-TL (grey surface, pdb code: 2YRQ). In the bar-graphs  $\alpha$ -helices are schematically represented on the top, missing residues are either prolines, or superimposed residues of the acidic IDR or absent because of exchange with the solvent, the dashed black line indicates the peak intensity decrease due to the titration dilution effect. Source data are provided as a Source Data file.

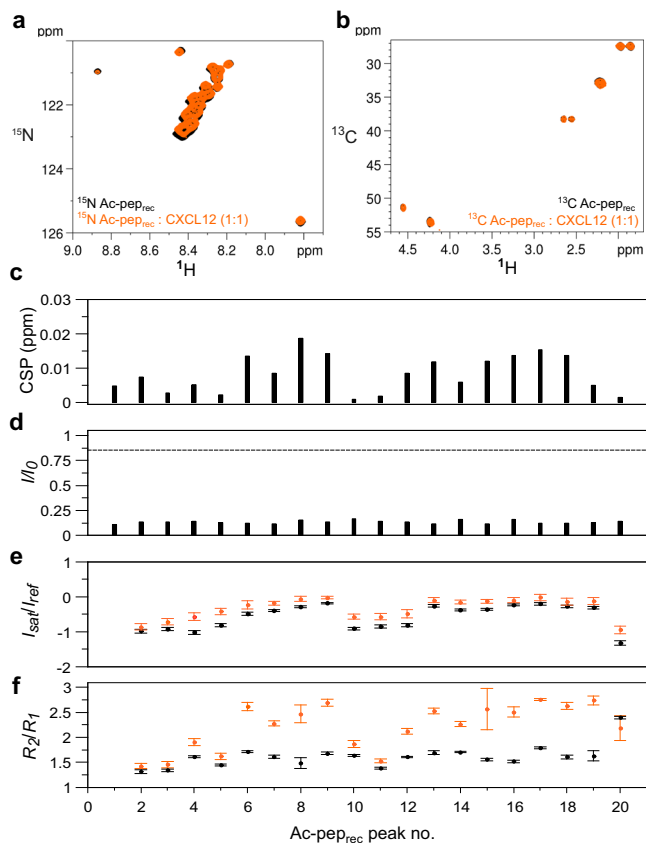
considered qualitative (Fig. 4c, d). Herein, analysis of  $^{15}\text{N}$  relaxation experiments (heteronuclear NOE,  $R_1$ ,  $R_2$ ) for free and bound  $^{15}\text{N}$  Ac-pep<sub>rec</sub> supported the formation of a fuzzy complex. Both states exhibited negative heteronuclear NOE values, signifying high flexibility on the picosecond to nanosecond timescale. A modest increase in complex heteronuclear NOE values indicated a slight reduction in peptide mobility in the presence of CXCL12. The increase in  $R_2/R_1$  and reduced peak intensity was in line with complex formation, with the consequent slowdown of its tumbling in solution and with the dynamic exchange between multiple CXCL12 binding sites (Fig. 4e, f). The presence of multiple interchangeable binding sites within the acidic IDR was further confirmed by titrating  $^{15}\text{N}$  CXCL12 with peptide fragments of the IDR (Ac-pep<sub>185-195</sub> and Ac-pep<sub>204-214</sub>), which similarly interacted with the CXCL12 dimerization surface (Supplementary Fig. 5a–f).

Taken together, the combination of the different NMR titrations and the relaxation data on  $^{15}\text{N}$  Ac-pep<sub>rec</sub> indicate that frHMGB1•CXCL12 is a fuzzy dynamic heterocomplex, characterized by multivalent inter- and intra-molecular equilibria involving CXCL12, the HMG tandem

domain and the acidic IDR as major players within this intricate network of interactions.

### The acidic IDR binds CXCL12 via long-range electrostatics

We next adopted the same dissection approach to investigate the thermodynamics of CXCL12 interaction with HMGB1, using a combination of isothermal titration calorimetry (ITC), microscale thermophoresis (MST) and fluorescence measurements. ITC injection of CXCL12 into Ac-pep solution generated spikes with a biphasic profile, indicative of different binding events, and large maximal exothermic heat changes ( $\sim -0.6 \mu\text{cal/s}$ ). Global fitting of the buffer-subtracted binding isotherm yielded an apparent  $K_{d1}$  of  $0.6 \pm 0.1 \mu\text{M}$  and  $K_{d2}$  of  $0.1 \pm 0.1 \mu\text{M}$ , in agreement with the low micromolar affinity estimated by NMR line-shape analysis (Fig. 5a, Table 1)<sup>38,39</sup>. Also the interaction between frHMGB1 and CXCL12 appeared biphasic and exothermic (Fig. 5b), though with one order of magnitude reduced amplitude ( $\sim -0.06 \mu\text{cal/s}$ ). The fitting of the curve yielded an apparent  $K_{d1} = 1.2 \pm 0.4 \mu\text{M}$  and a second one, whose nanomolar value should be taken with caution because of the large error in the global fitting<sup>39</sup>



**Fig. 4 | The acidic IDR peptide ( $^{15}\text{N}$  Ac-pep<sub>rec</sub>) dynamically interacts with CXCL12.** Superposition of (a)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$  Ac-pep<sub>rec</sub> (0.1 mM) without (black) and with (orange) CXCL12 (1:1), (b)  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectra of  $^{15}\text{N}/^{13}\text{C}$  Ac-pep<sub>rec</sub> (0.1 mM) without (black) and with (orange) CXCL12 (1:1). Bar graphs showing (c) chemical shift perturbation (CSPs) and (d) peak intensity ratios ( $I/I_0$ ) of  $^{15}\text{N}$  Ac-pep<sub>rec</sub> (0.1 mM) upon addition of CXCL12 (1:1). e Peak intensity ratios of heteronuclear NOE with and without proton saturation ( $I_{\text{sat}}/I_{\text{ref}}$ ), error bars were calculated by error propagation from the standard deviation (SD) of the average value of the noise in the saturated ( $\langle I_{\text{noise,sat}} \rangle$ ) and non saturated reference ( $\langle I_{\text{noise,ref}} \rangle$ ) spectra. f Ratios of  $R_2$  and  $R_1$  relaxation rates of  $^{15}\text{N}$  Ac-pep<sub>rec</sub> (0.1 mM) without (black) and with CXCL12 (1:1) (orange). Data of one representative experiment performed on  $n = 2$  biologically independent samples are presented as mean values  $\pm$  SD. Error bars are derived from relaxation data as described in methods. Amide resonances were not sequence specifically assigned and peaks were attributed arbitrary numbers. Source data are provided as a Source Data file.

(Table 1). Importantly, the low micromolar affinities measured by ITC for Ac-pep and frHMGB1 were in good agreement with the ones deriving from fluorescence and MST experiments, respectively (Fig. 5d, Table 1). While the heat of reaction between CXCL12 and frHMGB1-TL was not sufficient to derive any binding parameters (endothermic spikes at baseline level,  $-0.05 \mu\text{cal/s}$ ) (Fig. 5c), a change of the thermophoretic diffusion properties of fluorescently labeled CXCL12 was detectable in the presence of frHMGB1-TL (Fig. 5d). Notably, the derived affinity ( $K_d = 12.5 \pm 5.5 \mu\text{M}$ ) was almost one order of magnitude weaker than the one measured with full-length frHMGB1 ( $K_d = 1.7 \pm 0.2 \mu\text{M}$ ), thus confirming the important role of the acidic IDR in heterocomplex formation (Table 1).

We reasoned that the interaction might be dominated by long range electrostatic interactions between the acidic IDR and the basic surface of CXCL12 (Supplementary Fig. 6a, b). Indeed, the heat of reaction associated to CXCL12 interaction with Ac-pep or frHMGB1 at higher ionic strength (150 mM NaCl) did not yield any spikes above baseline (Supplementary Fig. 6c, d). Also in fluorescence and MST

experiments binding affinities were reduced by one order of magnitude, confirming the prominent contributions of long-range electrostatic interactions (Fig. 5e, Table 1). As further confirmation that the acidic IDR mainly interacts with CXCL12 via long range electrostatic interactions, we titrated into  $^{15}\text{N}$  CXCL12 a negative control peptide corresponding to the arginine/lysine-rich tail of the HMGB1 mutant form leading to the brachyphalangy, polydactyly and tibial aplasia malformation syndrome<sup>40</sup>. As expected, no interaction was detected (Supplementary Fig. 5g–i). However, as already mentioned, the acidic IDR is not the unique driving force in the interaction, as the oxidation status of BoxA is also relevant for the binding to CXCL12, as indicated by the reduced affinity of dsHMGB1 ( $K_d = 37.3 \pm 3.3 \mu\text{M}$ ) with respect to frHMGB1 (Supplementary Fig. 4e).

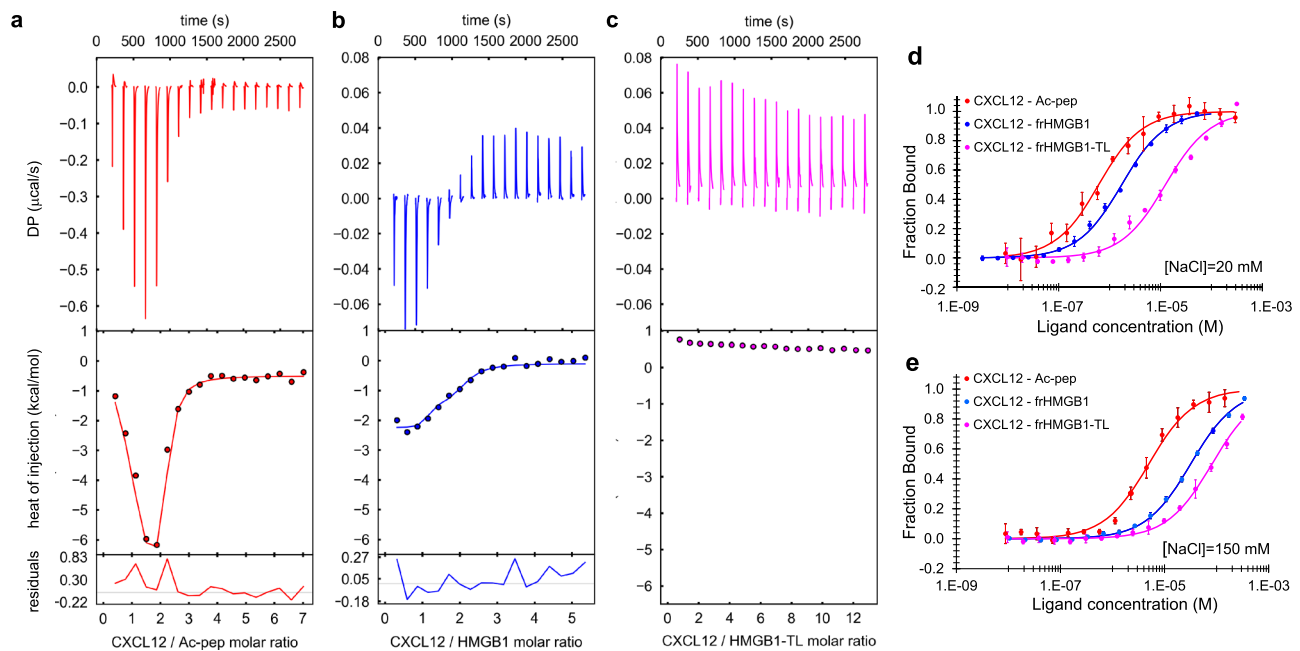
Collectively, ITC, MST, and fluorescence measurements, in accordance with NMR titrations, support a scenario in which the frHMGB1 tandem domain and the acidic IDR both contribute to heterocomplex formation, with the acidic IDR working as an antenna for recruiting CXCL12 via long-range electrostatic interactions.

### frHMGB1 and CXCL12 form an equimolar heterocomplex

Previously, we and others postulated that frHMGB1 and CXCL12 interact with a 1:2 ratio<sup>15,30</sup>. To verify this hypothesis, we rigorously assessed complex stoichiometry by analytical ultracentrifugation (AUC). Initially, we compared sedimentation velocity AUC (SV-AUC) experiments of the individual and combined components at different stoichiometric ratios<sup>41</sup>. Sedimentation coefficient distributions ( $c(s)$ ) of free CXCL12 (1.1 S) and frHMGB1 (2.3 S) were in good agreement with their apparent molecular weights (8.7 and 27 KDa, respectively) (Fig. 6a, b; Supplementary Table 1), and the corresponding frictional ratios ( $f/f_0$ ) of 1.3 and 1.4 were in line with the globular and elongated shape of CXCL12 and frHMGB1, respectively (Supplementary Table 1). At variance to stable complexes, characterized by distinct SV-AUC curves for the bound and individual components, the frHMGB1•CXCL12 heterocomplex presented only two separable sedimentation distributions: one relatively sharp peak at lower  $s$ , corresponding to free CXCL12, and one at higher  $s$  values, deriving from the sedimentation of free and bound frHMGB1. The latter, both at 50 mM (Fig. 6c, Supplementary Table 1) and 150 mM NaCl (Supplementary Fig. 7a–c, Supplementary Table 2), was relatively broad and shifted towards higher  $s$  values (and apparent MWs) with increasing CXCL12 concentrations. This behavior is typically observed in highly dynamic complexes, where the reaction boundaries between bound and unbound species cannot be resolved within the signal-to-noise ratio of the experiment<sup>42</sup>. Importantly, analysis of SV-AUC curves of frHMGB1-TL with increasing concentrations of CXCL12 showed only two peaks corresponding to the free components, and no sedimentation peaks displacement or broadening were observed. Conceivably, without the acidic IDR the association between the two components is too rapid to be detected in the sedimentation time scale (Supplementary Fig. 7d–f).

Next, to determine the ratio of frHMGB1 and CXCL12 in the complex distribution, we performed multi-signal sedimentation velocity analytical ultracentrifugation experiments (MSSV-AUC) and exploited the different extinction coefficients of frHMGB1 and CXCL12 at minimal (250 nm) and maximal (280 nm) wavelengths to distinguish between the two complex components<sup>43,44</sup>. Analysis of MSSV-AUC suggested that frHMGB1 predominantly forms with CXCL12 (or CXCL12-LM) a 1:1 heterocomplex in solution, as indicated by the equimolar concentrations for both CXCL12 (CXCL12-LM) and frHMGB1 obtained from the peak area at  $-3\text{S}$  (Fig. 7d–f, Supplementary Table 1–3, Supplementary Fig. 7g–k).

Based on these experiments we conclude that frHMGB1 and CXCL12, in contrast to previous assumptions, form an equimolar heterocomplex, with the acidic IDR playing a fundamental role in complex assembly.



**Fig. 5 | The acidic IDR of frHMGB1 interacts with CXCL12 via long-range electrostatic interactions.** ITC measurements of CXCL12 titrated into (a) Ac-pep (red), (b) frHMGB1 (blue) and (c) frHMGB1-TL (magenta) (20 mM TrisHCl at pH 7.5, 50 mM NaCl). The upper, middle and lower panels show, respectively, the ITC sequential heat pulses (DP, differential power) for binding, the integrated data corrected for heat of dilution and the residuals. Data in (a, b) were globally fitted. Data in c could not be fitted because the heat of reaction was too small to be fitted with a nonlinear least-squares method. Data represents peak integration of ITC

signal. One representative curve ( $n = 2$ ) for each titration is shown. Normalized variation of fluorescence of 5,6-FAM-labelled Ac-pep upon addition of CXCL12 (red) and normalized variation of MST signal of labeled CXCL12 in the presence of frHMGB1 (blue) and of frHMGB1-TL (magenta), with (d) 20 mM NaCl and with (e) 150 mM NaCl. Data are presented as mean values  $\pm$  SD of  $n = 3$  independent replicates. The thermodynamic parameters and  $K_d$ s are summarized in Table 1. Source data are provided as a Source Data file.

### SAXS supports the dynamic nature of frHMGB1•CXCL12

We then utilized small angle X-ray scattering (SAXS) to obtain low-resolution structural information on the estimated mass and shape of the heterocomplex. Primary analysis of the SAXS scattering curves of free CXCL12 and frHMGB1 yielded radii of gyration ( $R_g$ ) of 1.52 nm and 2.60 nm, respectively, as well as pairwise distance distribution plot,  $P(r)$ , derived maximum distance ( $D_{max}$ ) values of 4.8 nm and 8.5 nm, respectively. These values indicate that both proteins are monomeric and monodisperse in solution (Fig. 7a, Supplementary Table 4, Supplementary Fig. 8a, 9, 10). The values obtained for frHMGB1 are in accordance with a previous report<sup>27</sup>. The normalized Kratky representation of free CXCL12 and frHMGB1 displayed upward trends at higher  $qR_g$  values, well in agreement with the presence of flexible regions within a folded core (Fig. 7b). While CXCL12 displayed a relatively symmetric  $P(r)$  plot, the one of frHMGB1 had an asymmetric, elongated shape, tailed off to large distances (Supplementary Fig. 8b). This feature suggests the existence of multiple conformations in solution, and aligns with the presence of flexible regions and intramolecular multivalent fuzzy interactions between the acidic IDR and the basic HMG domains<sup>27</sup>. Thus, to generate a suitable ensemble of conformers able to account for HMGB1 intrinsic dynamics and match its experimental scattering curve, we used Ensemble Optimization Method (EOM)-based rigid body modelling<sup>27,45</sup> (Supplementary Fig. 11).

Analysis of the heterocomplex was challenging due to its transient and dynamic nature, which prevented its isolation through size exclusion chromatography and direct coupling to SAXS. Instead, we used a batch-mode strategy where frHMGB1 incubated with increasing CXCL12 concentrations was immediately analyzed by SAXS (Supplementary Fig. 12). In the SAXS spectrum of a protein complex the relative contribution of the individual components to the overall scattering pattern varies according to their size and scattering intensity<sup>46</sup>. Thus, the heterocomplex, which is the biggest component

(33.8 kDa), was expected to contribute most to the scattering signal and to dominate over the smallest one (CXCL12, 9.1 kDa). Indeed, in the presence of two equivalents of CXCL12, where more than 96% of HMGB1 was saturated with CXCL12, a 1:1 complex with an estimated molecular weight of 35 kDa was obtained (Supplementary Table 4). At higher CXCL12 concentrations, where a larger fraction of unbound CXCL12 was present, it was not possible to obtain a reliable estimate of the complex molecular weight (Supplementary Fig. 8a). We therefore focused our SAXS analysis on the sample comprising two CXCL12 equivalents. We observed an overall increase of the derived parameters with respect to free frHMGB1, compatible with complex formation ( $R_g$  of 2.93 nm and  $D_{max}$  of 10.5 nm) and a mass estimation corresponding to frHMGB1•CXCL12 heterocomplex (Supplementary Fig. 8a). The normalized Kratky plot and the  $P(r)$  plot confirmed the dynamic and flexible nature of the complex, containing both folded domains and unstructured segments (Fig. 7b, Supplementary Fig. 8b).

Next, to model more quantitatively the fuzzy nature of frHMGB1•CXCL12 heterocomplex, we combined NMR-guided rigid-body docking (SASREF<sup>47</sup> and FoXSDock<sup>48</sup>) to EOM<sup>45</sup>. Since our NMR data indicated that CXCL12 can interact with both BoxA and the acidic IDR, two initial docking models were generated. To guide the interaction between CXCL12 on BoxA we used SASREF<sup>47</sup> taking advantage of the experimental CSPs (Supplementary Fig. 13a). To model the association between CXCL12 and the acidic IDR, we opted for FoXSDock<sup>48</sup> (Supplementary Fig. 13b), which is more suitable when no prior information is available about the specific interacting regions within a disordered segment. Next, we used EOM to capture the intrinsic flexibility of frHMGB1 within the complex and to describe the size distribution of possible multiple conformations of frHMGB1•CXCL12. Herewith, we identified two EOM conformational ensembles fitting well the experimental data (with  $\chi^2 < 1$ , Supplementary Table 4). In the first CXCL12 binds to BoxA of frHMGB1, adopting alternatively open or more collapsed conformations (Fig. 7c). In the second one, the acidic

**Table 1 | Thermodynamic parameters of the interactions between CXCL12 and Ac-pep, frHMGB1 and frHMGB1-TL**

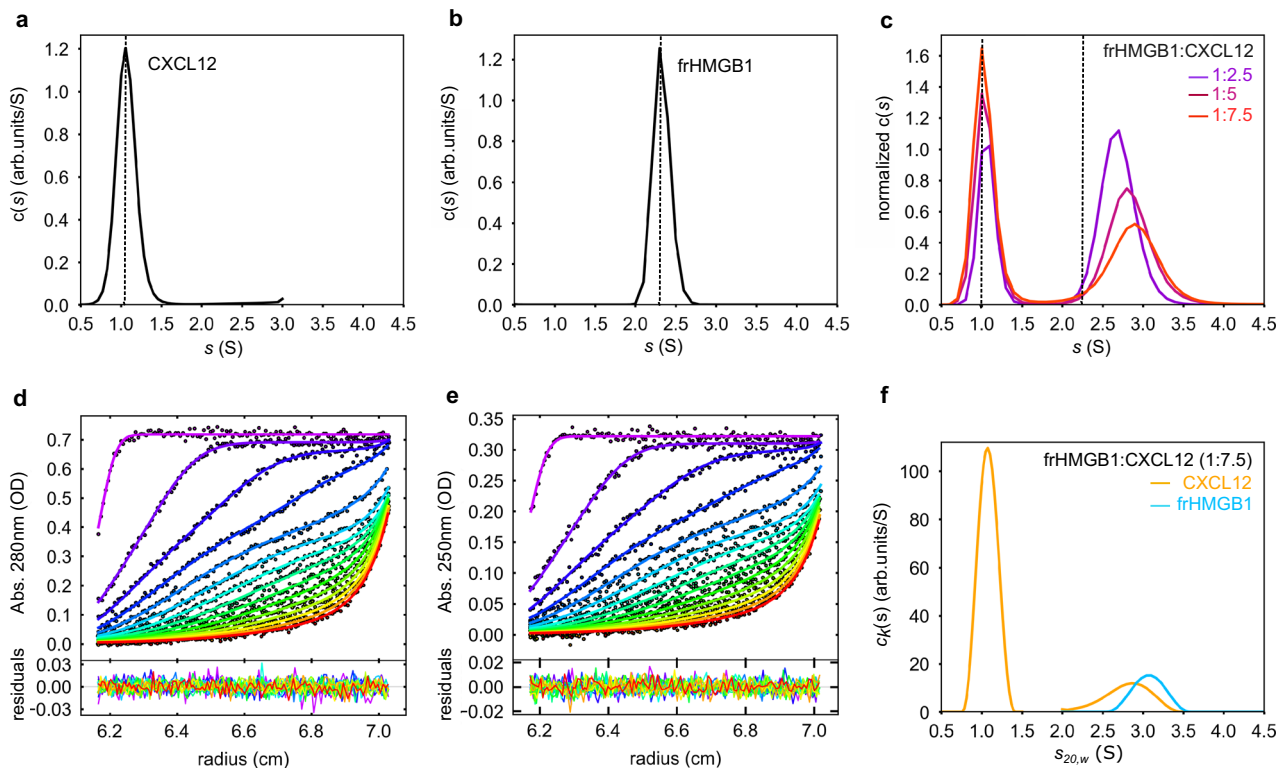
	ITC <sup>a</sup>				MST	
	$Kd_1$ ( $\mu\text{M}$ )	$Kd_2$ ( $\mu\text{M}$ )	$\Delta H_1$ (kcal/mol)	$\Delta H_2$ (kcal/mol)	$Kd_{[\text{NaCl}]=20\text{mM}}$ ( $\mu\text{M}$ )	$Kd_{[\text{NaCl}]=150\text{mM}}$ ( $\mu\text{M}$ )
Ac-pep	0.6 ± 0.1	0.1 ± 0.1	-10.4 ± 0.8	1.6 ± 0.7	0.6 ± 0.1 <sup>b</sup>	5.0 ± 0.6 <sup>b</sup>
frHMGB1	1.2 ± 0.4	$7.8 \times 10^{-3} \pm 1.5 \times 10^{-2}$	-1.4 ± 0.1	-2.1 ± 0.2	1.7 ± 0.2 <sup>c</sup>	31.8 ± 1.4 <sup>c</sup>
frHMGB1-TL	n.d.	n.d.	n.d.	n.d.	12.5 ± 5.5 <sup>c</sup>	81.4 ± 6.8 <sup>c</sup>

<sup>a</sup>CXCL12 is the titrant,  $n=2$  independent replicates, error estimates from covariance matrix.

<sup>b</sup>CXCL12 is the titrant,  $n=3$  independent replicates, values are mean ± standard deviation.

<sup>c</sup>frHMGB1 or frHMGB1-TL is the titrant,  $n=3$  independent replicates, values are mean ± standard deviation.

n.d. not detected.



**Fig. 6 | frHMGB1 and CXCL12 form a transient 1:1 heterocomplex.** Sedimentation coefficient distributions  $c(s)$  of (a) free CXCL12 (38.2  $\mu\text{M}$ ) and (b) frHMGB1 (15.6  $\mu\text{M}$ ), scanned by absorbance at 280 nm. c Overlay of normalized  $c(s)$  showing the interaction between frHMGB1 (7.8  $\mu\text{M}$ ) and increasing concentrations of CXCL12 (colors). The dotted lines indicate the sedimentation coefficients of the free components. d, e Global multi-signal sedimentation velocity analysis to determine the stoichiometry of frHMGB1: CXCL12 complex, with 7.7  $\mu\text{M}$  frHMGB1 and 43  $\mu\text{M}$  CXCL12. The raw sedimentation signals of frHMGB1: CXCL12 mixture

acquired at different time points with (d) absorbance at 280 nm, and (e) absorbance at 250 nm with the corresponding signal profiles as a function of radius in centimeters. The time-points of the boundaries are indicated in rainbow colors, progressing from purple (early scans) to red (late scans). Only every 3rd scan used in the analysis are shown. Residuals of the fit are shown at the bottom. f Decomposition into the component (k) sedimentation coefficient distributions,  $c_k(s)$ , for CXCL12 (yellow line) and frHMGB1 (cyan line). Source data are provided as a Source Data file.

IDR of frHMGB1 interacts with CXCL12, while the two HMG boxes adopt different reciprocal orientations (Fig. 7d).

Overall, SAXS analysis supports the notion that frHMGB1 and CXCL12 form an equimolar fuzzy complex.

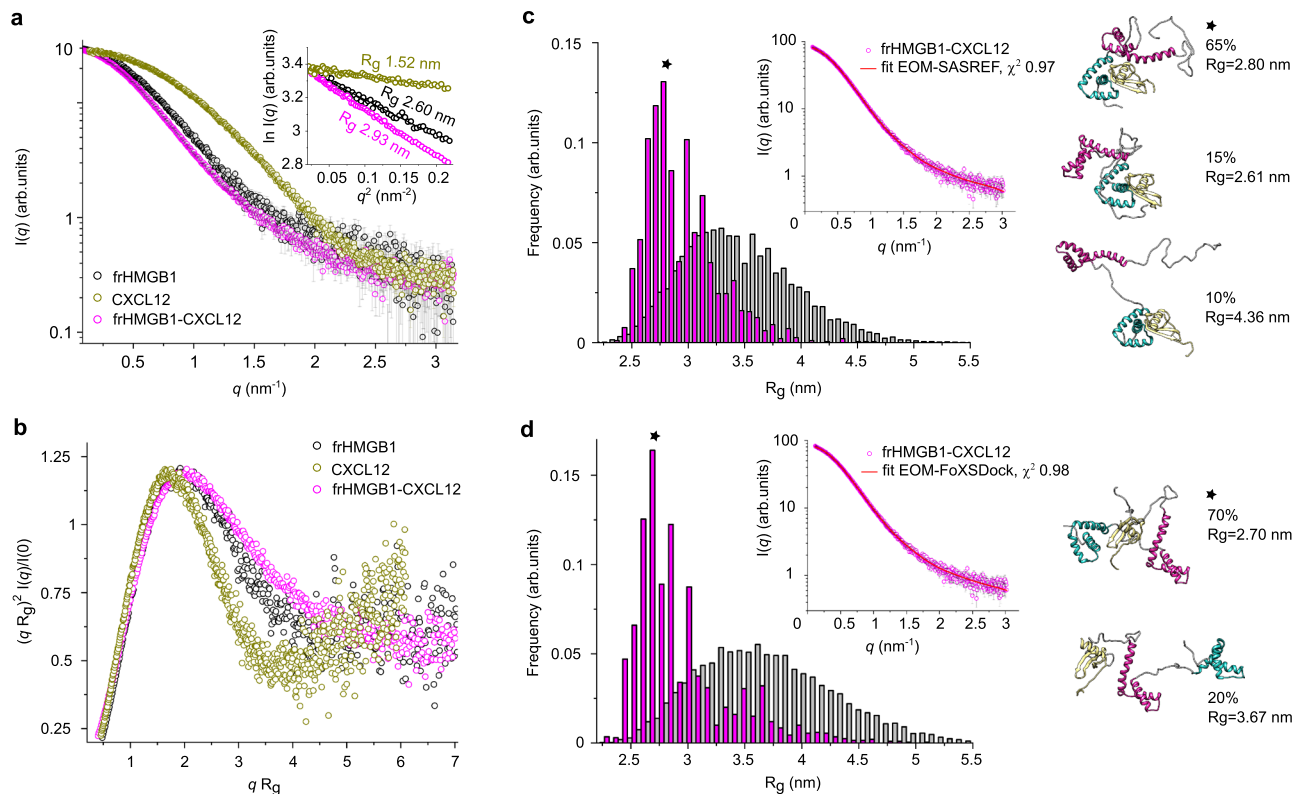
**The acidic IDR modulates frHMGB1•CXCL12 binding to CXCR4**  
Having shown that the acidic IDR of frHMGB1 plays a major role in the interaction with CXCL12, we tested whether it also played a role in the binding of the heterocomplex to its CXCR4 receptor. For these experiments, we chose mouse AB1 cells, a cellular model of malignant mesothelioma<sup>49</sup>, as they express high levels of CXCR4<sup>23</sup>.

We first verified the existence of the frHMGB1•CXCL12 heterocomplex in association with CXCR4 on the cell membrane. Proximity ligation assays (PLA) between HMGB1 and CXCL12 clearly identified the frHMGB1•CXCL12 heterocomplex on the surface of AB1 cells, and the

PLA signal was competed by increasing concentrations of AMD3100, a specific CXCR4 antagonist (Fig. 8a). To confirm that the heterocomplex binds CXCR4 we deleted both *Cxcr4* alleles in AB1 cells by CRISPR/Cas9 mutagenesis. As expected, PLA on the surface of AB1 *Cxcr4* KO cells detected only background levels of the HMGB1•CXCL12 heterocomplex (Fig. 8b). Then, we treated wildtype AB1 cells with increasing amounts of Ac-pep, which significantly decreased the amount of detectable heterocomplex, in line with the NMR observation that Ac-pep can bind CXCL12 and disrupts the frHMGB1•CXCL12 heterocomplex.

We then tested the binding to cell surface CXCR4 of heterocomplexes made with full-length or tailless frHMGB1 (frHMGB1-TL). Cells were washed at acidic pH to remove ligands bound to receptors, and then exposed to preformed mixtures of CXCL12 and full-length or frHMGB1-TL; cells were kept at 4 °C to prevent receptor





**Fig. 7 | SAXS studies of free frHMGB1, CXCL12 and of frHMGB1-CXCL12.** **a**  $I(q)$  versus  $q$  experimental SAXS profiles for CXCL12 (gold), frHMGB1 (black) and frHMGB1-CXCL12 complex (magenta). The curves are shifted by an arbitrary offset for better comparison. Error bars represent an estimate of the experimental error  $\sigma$  on the intensity recorded for each value of  $q$  as assigned by data reduction software. In the inset, the Guinier regions used to estimate the radii of gyration ( $R_g$ , nm). **b** Dimensionless Kratky plots for the data presented in (a). Analysis of SAXS data by EOM on frHMGB1-CXCL12 models obtained using (c) SASREF and (d) FoXSdock with distributions of the selected ensemble conformers (magenta bars) and the

initial pools of structures (grey bars) as a function of  $R_g$  in nm. In the insets,  $I(q)$  versus  $q$  (magenta squares) with the EOM fitting (red lines with the corresponding  $\chi^2$  values) for the frHMGB1-CXCL12 complex. Representative structures of the most populated EOM ensembles are shown in cartoon, with BoxA, BoxB, acidic IDR and CXCL12 coloured in cyan, magenta, grey and gold, respectively. For each ensemble, the frequency-weighted size average (the asterisks indicate the most populated fractions) and  $R_g$  values are indicated. Source data are provided as a Source Data file.

internalization. Both full-length and tailless frHMGB1 formed complexes that could be detected by PLA on the cell surface; however, frHMGB1-TL formed fewer complexes and did not reach saturation (Fig. 8c).

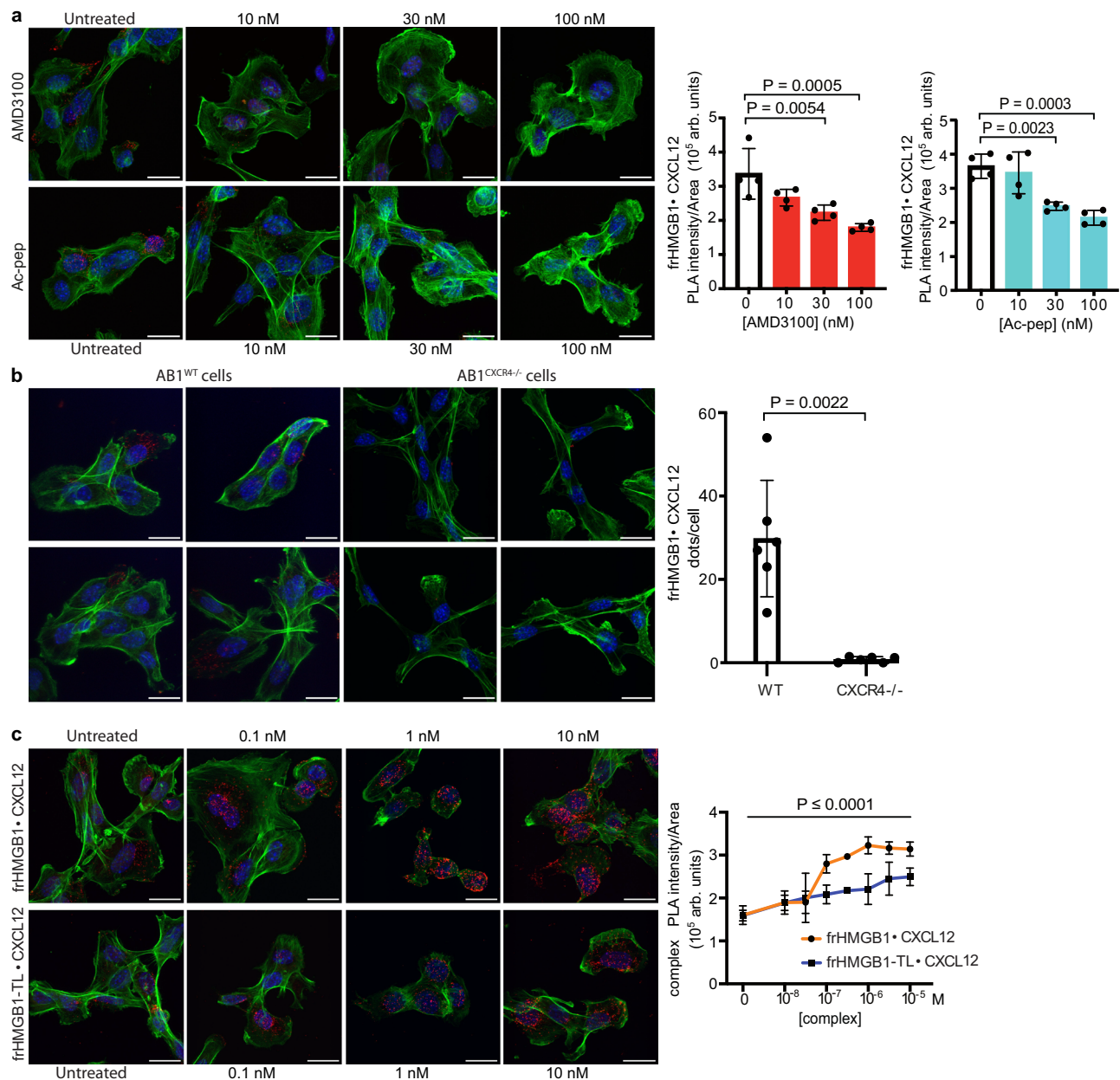
Together, these results show that the frHMGB1 acidic IDR facilitates the binding of the frHMGB1-CXCL12 heterocomplex to its CXCR4 receptor, and that the Ac-pep competes with the acidic IDR for heterocomplex formation and binding to CXCR4.

## Discussion

Until now, structural studies aimed at getting mechanistic insights into frHMGB1-CXCL12 have focused on the interaction of CXCL12 with the single isolated HMG boxes<sup>15,30</sup>. In these studies the authors took advantage of NMR CSPs obtained from titrations of CXCL12 with the isolated HMG boxes to produce a 3D model in which each single HMG box bound one CXCL12 molecule<sup>15,30</sup>. The question arose whether such a simplified approach, in which the HMG boxes are separated and the IDRs (i.e. the acidic C-terminal tail and the linker connecting the HMG boxes) are neglected, can faithfully recapitulate the interaction of CXCL12 with the full-length protein. IDRs within their host proteins are often the major players in the recognition of their partners<sup>50-52</sup>, and they are generally characterized by low complexity sequences, often containing charged residue<sup>53</sup>. In several cases IDRs maintain a significant level of disorder in their bound states, exchanging between multiple conformations and giving rise to so-called fuzzy complexes<sup>31,32,37,54</sup>. Their structural plasticity and ability to be engaged in fuzzy interactions renders IDRs well-suited to favour the formation

of biomolecular condensates through liquid-liquid phase separation, where specific interactions can be established without a clearly defined bound-state conformation<sup>32</sup>. Moreover, their conformational malleability allows IDRs to engage with various partners and in different cellular conditions, thus broadening the interaction scopes of their host protein<sup>35</sup>. This also applies for HMGB1, where the acidic C-terminal IDR modulates interactions with nucleic acids<sup>56</sup> and different proteins, such as histones<sup>57,58</sup> and p53<sup>59</sup>. Importantly, replacement of the acidic IDR with an arginine-rich basic tail has been recently shown to cause a complex human malformation syndrome, which results from HMGB1 aberrant phase separation in the nucleolus and nucleolar dysfunction<sup>40</sup>.

In accordance with its functional relevance, here we show that the acidic IDR, whether isolated or in the context of the full-length protein, binds to the CXCL12 homodimerization surface. Interestingly, the CXCL12 residues involved in the interaction with both Ac-pep and frHMGB1 in part coincide with the ones involved in the interaction with negatively charged heparin oligosaccharides<sup>60</sup>. Importantly, the acidic IDR, because of its repetitive sequence, contains multivalent sites able to interact with CXCL12, as indicated by NMR titrations of <sup>15</sup>N CXCL12 with synthetic peptide fragments of the acidic IDR. In line with this multivalency, NMR relaxation experiments performed on the <sup>15</sup>N Ac-pep<sub>rec</sub> show that the acidic IDR maintains its conformational flexibility also when bound to CXCL12. Noteworthy, binding of full-length frHMGB1 to CXCL12 is significantly weaker than to the isolated acidic IDR, indicating competition of the intra-molecular frHMGB1 boxes with CXCL12 for the acidic IDR. Both ITC and MST experiments suggest



**Fig. 8 | The acidic IDR modulates frHMGB1-CXCL12 binding to CXCR4 on AB1 cells.** **a** Representative confocal microscopy images of Proximity Ligation Assays (PLAs) performed on the frHMGB1-CXCL12 complex on the surface of AB1 malignant mesothelioma cells. Cells were either untreated or treated with AMD3100 or Ac-pep (light blue bars) for 1 h at 37 °C/5% CO<sub>2</sub> at three different concentrations (10, 30, or 100 nM). PLA signal was quantified as described in the Methods section. Mean ± SD are indicated;  $n = 4$  FOV (field of view) per concentration. One-way ANOVA was performed comparing AMD3100 (red) and Ac-pep (light blue) treated cells to untreated cells (white); AMD3100 versus untreated:  $**P = 0.0054$ ;  $***P = 0.0005$ ; Ac-pep versus untreated:  $**P = 0.0023$ ,  $***P = 0.0003$ . Data are presented as arbitrary units (arb. units). Scale bar; 20  $\mu$ m. **b** Representative confocal microscopy images of PLAs performed on the frHMGB1-CXCL12 complex on the surface of either wild type (WT) or *Cxcr4* knockout (*Cxcr4*<sup>-/-</sup>) AB1 malignant mesothelioma cells. PLA signal was quantified as number of dots per cell in FOV.

Statistical analysis; two-tailed, non-parametric Mann-Whitney test. Mean ± SD are indicated;  $n = 6$  FOV per condition. Data are presented as arbitrary units (arb. units). Scale bar; 20  $\mu$ m. **c** PLA signal quantification on the surface of ligand-stripped AB1 cells exposed at 4 °C to increasing concentrations of either frHMGB1-CXCL12 or frHMGB1-TL-CXCL12 equimolar heterocomplexes. Data are presented as arbitrary units (arb. units). Mean ± SD are indicated;  $n = 4$  FOV per concentration. The difference between frHMGB1-CXCL12 (orange) and frHMGB1-TL-CXCL12 (blue) heterocomplexes is statistically significant ( $P < 0.0001$ ) by two-way ANOVA. Sidak's multiple comparisons test revealed statistically significant differences between frHMGB1-CXCL12 and frHMGB1-TL-CXCL12 at the following concentrations; 10<sup>-7</sup> M:  $P = 0.0004$ , 10<sup>-6.5</sup> M,  $P = 0.0003$ , 10<sup>-6</sup> M:  $P < 0.0001$ , 10<sup>-5.5</sup> M:  $P = 0.0013$ , and 10<sup>-5</sup> M:  $P = 0.0032$ . In all panels, nuclei are in blue (Hoechst 33342), phalloidin is in green, and the frHMGB1-CXCL12 PLA signal is red. Scale bar; 20  $\mu$ m. Source data are provided as a Source Data file.

long range electrostatic interactions between the acidic IDR and the basic surface of CXCL12 as major drivers for binding, as increased ionic strength or deletion of the acidic IDR reduced binding to CXCL12. Long-range electrostatic interactions, though fundamental, are not the unique driving force for complex formation, as both NMR and MST indicate that tailless frHMGB1 (frHMGB1-TL) binds with micromolar affinity to the dimerization surface of CXCL12, albeit less well with

respect to the full-length protein. Thus, structured and unstructured frHMGB1 regions potentially recognize the same CXCL12 surface, which behaves as a structural hub for multivalent interactions. Notably, while previous NMR titrations suggested that the isolated HMG-boxes interact similarly with the dimerization surface of CXCL12<sup>15</sup>, a preference for BoxA is apparent in the context of the tandem HMG-boxes, with the interaction surface mainly involving the two short

helices of the HMG domain. Of note, the targeting of BoxA is in line with the ability of CXCL12 to preferentially recognize the reduced forms of Cys-22 and Cys-44<sup>20</sup>. Moreover, our MST and NMR experiments show that, while dsHMGB1 is still able to bind to CXCL12, most likely via the acidic IDR, its affinity is starkly lower, thus confirming the important contribution of BoxA to the interaction. Privileged binding to BoxA is also in agreement with the equimolar stoichiometry of the heterocomplex suggested by both AUC and SAXS experiments.

Collectively, our data reveal that the frHMGB1•CXCL12 heterocomplex behaves as a typical fuzzy complex<sup>31</sup>, whose formation relies on an intricate network of inter- and intra-molecular interactions of comparable affinities. As commonly observed in fuzzy binding, at least one of the elements, in this case the acidic IDR, is dynamic, maintains its conformational flexibility in the bound state and is fundamental for the interaction<sup>36,61</sup>. In addition, the intrinsic independent rotation of the two HMG-boxes provides an additional dynamic level to the system. Thus, the heterocomplex cannot be described by a unique structure, but is best represented by a heterogeneous ensemble of structures reflecting the different ongoing equilibria. Accordingly, the SAXS data of the heterocomplex are best fit by different plausible docking models obtained by EOM, where CXCL12 binds frHMGB1 in a promiscuous manner, alternatively associating to the manifold conformations of the acidic IDR and the different BoxA orientations.

Based on our data we propose a model (Fig. 9) in which the acidic IDR works as a wrapping antenna that recruits CXCL12 through long-range electrostatic interactions. Being intrinsically disordered, the acidic IDR does not present a single binding site to CXCL12 but rather resembles a diffuse binding cloud, in which multiple nearly-identical binding sites are dynamically distributed<sup>62,63</sup>, preserving a significant flexibility even in bound states. This behavior is reminiscent of the mean electrostatic field created by multiple phosphates between the disordered yeast cyclin-dependent kinase (CDK) inhibitor Sic1 and its cognate binding partner, the F-box protein Cdc4<sup>36</sup>. The rapid on-off rate usually associated to long-range electrostatic interactions allows then CXCL12 to interact with HMG-boxes, in particular with BoxA in its reduced form, partially outcompeting HMGB1's intramolecular contacts. Remarkably, the binding mode of frHMGB1 to CXCL12 radically differs from the usual beta-beta or alpha-beta interactions observed in chemokine heterocomplexes<sup>4</sup>.

CXCL12 bound to frHMGB1 can then accommodate inside the cradle formed by the transmembrane helices of CXCR4, yielding a three-component complex (Fig. 9), whose existence on the cell surface is supported by PLA experiments. These results are in line with the observation that the HMGB1•CXCL12 heterocomplex acts differentially from CXCL12 alone on the reorganization of the actin cytoskeleton and on  $\beta$ -arrestin recruitment<sup>22</sup>. Both actions depend on CXCR4. So far, however, our data do not specify whether a specific subset of the conformations of the HMGB1•CXCL12 heterocomplex preferentially binds to CXCR4. In our cellular context, the acidic IDR facilitates the binding of the frHMGB1•CXCL12 heterocomplex to CXCR4, and Ac-Pep does compete with it, confirming the results obtained by NMR, and implying that the acidic IDR plays a major role in the formation of the CXCR4•frHMGB1•CXCL12 ternary complex as well. The fuzziness of the interactions in the frHMGB1•CXCL12 heterocomplex forces us to reconsider the structure of HMGB1, which itself can be considered fuzzy: in solution HMGB1 populates an ensemble of different microstates in which the D/E repeats are associated through electrostatic transient interactions with different segments of HMGB1, that in turn are partially screened and only transiently exposed to natural interactors<sup>26</sup>. Plausibly, the fuzzy conformation of HMGB1 allows interactions with multiple partners<sup>64</sup>, all of which have micromolar-range apparent affinities. Indeed, HMGB1 often works as a chaperone, by binding one interactor and facilitating its further interaction with another molecule<sup>65–67</sup>. As such, the mechanism and the conformational

heterogeneity through which HMGB1 binds to CXCL12 is in part reminiscent of other chaperones, like small heat shock proteins, that exploit their large IDR to transiently interact with their clients<sup>68</sup>.

In brief, we show here that frHMGB1 and CXCL12 form a bimolecular heterocomplex, which is fuzzy and highly dynamic, and can go on to form a ternary complex with the CXCR4 receptor. These conclusions are in line with the concepts of HMGB1 working as a molecular chaperone and of IDPs/IDRs being typically involved in signaling, regulation, recognition, and control of various cellular pathways<sup>69</sup>.

Finally, recent studies have revealed that inhibiting heterophilic interactions among chemokines interactions can reduce inflammation<sup>13</sup>. Consequently, targeting these protein-protein interactions is emerging as a valuable strategy for developing selective antagonists to finely modulate specific inflammatory responses. This holds true also for the frHMGB1•CXCL12 heterocomplex, a key regulator of inflammatory cell recruitment via the CXCR4 axis, making it an attractive target for selective anti-inflammatory agents. Indeed previous research has demonstrated that HMGB1•CXCL12 is druggable exploiting specific pockets within HMGB1 and CXCL12<sup>24</sup>. As the targeting of IDR and IDPs is starting to emerge as a pharmacological strategy<sup>70</sup>, we anticipate that interfering with the fuzzy interactions within the HMGB1•CXCL12 heterocomplex could represent an additional opportunity to inhibit its detrimental activity in inflammatory conditions.

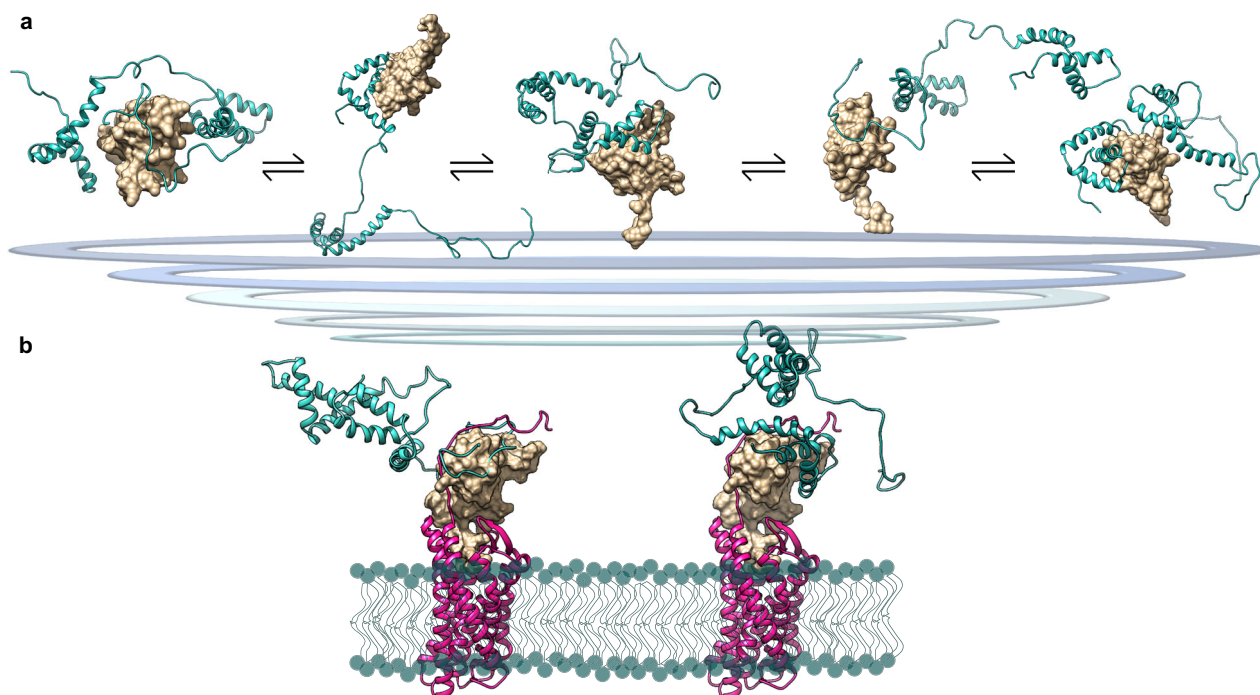
## Methods

### Protein production and synthetic peptide

Recombinant labeled and unlabelled (<sup>15</sup>N) HMGB1 (uniprot code P63159, residues 2–215, renumbered from 1 to 214) and HMGB1-TL (uniprot code P63159, residues 2–188, renumbered from 1 to 187) were transformed in BL21 (DE3) pLysS and BL21 (DE3) strains of *E. coli*, respectively, using the pETM-II expression vector (EMBL, Heidelberg, DE). The proteins were purified as described in Supplementary Information. dsHMGB1 and dsHMGB1-TL were obtained by extensively dialyzing their fully reduced forms against a buffer devoid of Dithiothreitol (DTT, volume ratio sample:buffer 1:2000), changing the buffer every 2 h overday, followed by an overnight dialysis. To reach the fully oxidized state, the samples were incubated at 4 °C for 72 h, before NMR and MST analysis. The oxidation status of HMGB1 constructs was checked by <sup>1</sup>H-1D NMR monitoring the aliphatic spectral region, which displays distinctive features for the oxidized and reduced forms. Recombinant labeled and unlabelled <sup>15</sup>N/<sup>13</sup>C CXCL12 (uniprot code P48061, residues 23–89, renumbered from 1–68) and CXCL12-LM were transformed into *E. coli* strain BL21 (DE3) using the expression vector pET30a. The proteins were purified as described in Supplementary Information. Unlabeled N-terminal-6His-tagged CXCL12 for MST measurements was provided by HMGBiotech (Milan, Italy).

For the production of CXCL12-LM, site-directed mutagenesis was performed to introduce mutations L55C and I58C in CXCL12 pET30a expression vector by using standard overlap extension methods (Supplementary Information). The DNA constructs were sequenced by Eurofins (Milan, Italy). Protein concentrations were determined considering molar extinction coefficients at 280 nm of 21430, 8730 M<sup>-1</sup>cm<sup>-1</sup>, 8855 M<sup>-1</sup>cm<sup>-1</sup> for HMGB1 (and HMGB1-TL), CXCL12 and CXCL12-LM, respectively.

Recombinant Ac-pep (Ac-pep<sub>rec</sub>) (corresponding to the HMGB1 acidic IDR, 30 aa, uniprot code P63159, residues 186–215, renumbered from 185 to 214) was produced as 6His-SUMO3 tagged fusion protein using the pETM11-SUMO3 expression vector (EMBL, Heidelberg, DE). Cloning was performed by GENEWITZ from Azenta Life Sciences. Expression was carried out in BL21 (DE3) *E. coli* cells by induction at 30 °C for 18 h with 0.5 mM Isopropil- $\beta$ -D-1-thiogalattopyranoside (IPTG). Cells were harvested by centrifugation (8200  $\times$  g, 4 °C) and lysed by sonication in Lysis Buffer (150 mM NaCl, 20 mM Tris-HCl pH 8, 0.2% NP-40, 2 mM  $\beta$ -mercaptoethanol (BME), Complete EDTA-free



**Fig. 9 | Model of frHMGB1-CXCL12 fuzzy complex.** **a** Representative EOM models of the fuzzy interactions between frHMGB1 (cyan cartoon) and CXCL12 (gold surface). **b** Explicative representations of possible different frHMGB1-CXCL12 conformations bound to CXCR4. Two SAXS-EOM frHMGB1-CXCL12 models have been

superimposed on the theoretical model of CXCL12 in complex with CXCR4<sup>85</sup>, with CXCR4 in magenta, CXCL12 in gold surface and frHMGB1 in cyan; the lipid bilayer is represented with cyan spheres and lines.

protease inhibitor (Roche), 2 µg/mL DNase, 20 µg/mL RNase). The 6His-SUMO3-tagged protein was purified on a histidine affinity column (Ni-NTA agarose, Quiagen) and eluted with Elution Buffer (150 mM NaCl, 20 mM Tris-HCl pH 8, 2 mM BME, 300 mM imidazole). The 6His-SUMO3 tag was cleaved during overnight dialysis at 4 °C against Dialysis Buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM BME, 5% glycerol) by addition of 6His-tagged Small Ubiquitin-Related Modifier (SUMO)-Specific Protease 2 protease (SEN2, homemade). 6His-SUMO3 tag and 6His-SEN2 were then separated from the digested Ac-pep<sub>rec</sub> by a second purification step on a Ni-NTA column. The peptide (having a serine at the N-terminus, deriving from cloning) was further purified on an HiTrap Q HP anion exchange chromatography column (Cytiva) using buffer A (20 mM Tris-HCl pH 8, 50 mM NaCl) and a slow linear gradient (3 ml/min flow, 15 column volumes) of buffer B (20 mM Tris-HCl pH 8, 1 M NaCl). The purified peptide was dialyzed against Milli-Q, lyophilized and resuspended into NMR buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 6, 20 or 150 mM NaCl). The peptide identity was confirmed by mass spectroscopy. Peptide concentration was determined using the absorbance at 205 nm and the molar extinction coefficient at 205 nm (83,400 M<sup>-1</sup> cm<sup>-1</sup>) calculated by the web server [<http://nickanthis.com/tools/a205.html>]. Uniformly labeled <sup>15</sup>N and <sup>15</sup>N/<sup>13</sup>C Ac-pep<sub>rec</sub> was produced by growing BL21 (DE3) *E. coli* cells in M9 minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl, with or without <sup>13</sup>C D-glucose.

Synthetic Ac-pep (corresponding to HMGB1 acidic IDR, 30 aa, uniprot code P63159, residues 186–215, renumbered from 185 to 214) was purchased from Caslo Lyngby, Denmark. Synthetic Ac-pep fragments corresponding to residues 185–195 (Ac-pep<sub>185-195</sub>) and residues 204–214 (Ac-pep<sub>204-214</sub>) and the synthetic peptide corresponding to the arginine rich tail present in an aberrant HMGB1 form<sup>40</sup> (R<sub>185</sub>RKMRKMKRMRRRRKMKKMKMKMKMKMNK<sub>214</sub>, whereby M<sub>210</sub>-M<sub>211</sub> where mutated into lysines to avoid possible aggregation problems deriving from four consecutive methionines) were purchased from Davids Biotechnologie GmbH (Germany). Peptide purity

(>95%) was confirmed by HPLC and mass spectrometry. Peptide concentration was estimated from its dry weight. For ITC measurements with Ac-pep, to obtain a more accurate estimation of the concentration by UV a tyrosine ( $\epsilon_{274nm} = 1405 \text{ M}^{-1} \text{ cm}^{-1}$ ) was added at the peptide N-terminus, for MST experiments 5,6 FAM (5(6)-Carboxyfluorescein) was added at the peptide N-terminus.

#### NMR spectroscopy (NMR)

NMR experiments were performed at 298 K on a Bruker Avance 600 MHz equipped with inverse triple-resonance cryoprobe and pulsed field gradients (Bruker, Karlsruhe, Germany). Typical samples concentration was 0.1–0.4 mM. Data were processed using NMRPipe 10.9<sup>71</sup> or Topspin 3.26 (Bruker) and analyzed with CCPNmr Analysis 2.4<sup>72</sup>. The <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N chemical shifts of CXCL12 in the presence of Ac-pep and of CXCL12-LM were obtained from three-dimensional HNCA, CBCA(CO)NH, CBCANH, HNC0 experiments (Supplementary Fig. 14, 96% amides assigned).

#### NMR titration experiments

Before NMR titrations the samples (titrant and titrated solution) were dialyzed against the same buffer, 20 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 6.3, 20 mM NaCl, supplied with 0.15 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and D<sub>2</sub>O (10% v/v). Of note, HMGB1 upon removal of DTT maintained its thiol form for at least 12 h. In the case of titrations with synthetic peptides the lyophilized peptides were dissolved directly in the NMR buffer, where necessary the pH was adjusted with 0.1 M NaOH. Titrations were carried out by adding to <sup>15</sup>N labelled protein samples (typically 0.1 mM) small aliquots of concentrated (15 mM) peptide stock solutions or unlabelled protein (0.6–1 mM). For each titration point (0.5, 1, 1.5, 2 equivalents of ligand) a 2D water-flip-back <sup>15</sup>N-edited HSQC spectrum was acquired with 2048 (160) complex points, apodized by 90° shifted squared (sine) window functions and zero filled to 2048 (512) points for <sup>1</sup>H (<sup>15</sup>N). Spectra assignment was made following individual cross-peaks through the titration series. For

each residue the weighted average of the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift perturbation (CSP) was calculated as

$$\text{CSP} = \sqrt{\frac{\Delta\delta\text{H}^2 + \Delta\delta\text{N}^2}{25}} \quad (1)$$

where  $\Delta\delta\text{H}$  and  $\Delta\delta\text{N}$  are, respectively, the differences of  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts between free and bound protein<sup>73</sup>. The corrected standard deviation ( $\sigma_0$ ) was calculated as described in<sup>74</sup>. Because of extensive line broadening due to ligand binding in the intermediate exchange regime on the NMR time scale, we also monitored changes in the intensity ratio ( $I/I_0$ ) of the  $^1\text{H}$ - $^{15}\text{N}$  amide resonances, where  $I_0$  and  $I$  are the peak intensities in the free and bound protein, respectively.

### Lineshape analysis

We performed 2D NMR lineshape analysis using the software TITAN 1.6<sup>33</sup>. Spectra were processed with NMRpipe 10.9<sup>71</sup> with a script provided by TITAN 1.6. A series of regions of interest containing isolated peaks with CSPs  $> \text{avg} + \text{SD}$  were selected (V18, V23, K24, H17, A40, H25 and R12 for  $^{15}\text{N}$ -CXCL12 and V23, K24, H25 for  $^{15}\text{N}$ -CXCL12\_LM) (Supplementary Fig. 2b) and fitted by optimizing the chemical shifts and line widths for the free and the bound state. The program estimates  $K_D$ ,  $k_{\text{off}}$  (we fixed 1:1 stoichiometry). Error estimates for the fit parameters were obtained using the bootstrap resampling of residuals procedure implemented in TITAN 1.6<sup>33</sup>.

### NMR relaxation measurements

Measurements of heteronuclear  $\{^1\text{H}\}$ - $^{15}\text{N}$  nuclear Overhauser (NOE) enhancement, longitudinal and transversal  $^{15}\text{N}$  relaxation rates ( $R_1$ ,  $R_2$ ) on free  $^{15}\text{N}$  Ac-pep (0.3 mM) and in the presence of CXCL12 (1:1) were performed using standard  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra with varying relaxation delays<sup>75</sup>. Duty-cycle heating compensation were used for both  $T_1$  and  $T_2$  relaxation experiments<sup>76</sup>.  $T_1$  and  $T_2$  decay curves were sampled at 10 and 9 different relaxation delays, respectively and two duplicate delays (ms). ( $T_1$ : 50/50, 150, 250/250, 450, 650, 900, 1100, 1400, 2000, 3000 and  $T_2$ : 72, 12/12, 28, 44, 56, 112/112, 144, 200, 244) collected in random order, with 4 s recovery delay. Two delays were acquired twice for evaluation of the average ( $T_1$ ), ( $T_2$ ) values and the corresponding standard deviations. The  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOEs were measured recording HSQC spectra with and without proton saturation in an interleaved fashion using a 6 s recycle delay/saturation. The standard deviation of the noise of both saturated and unsaturated spectra were used to estimate  $I_{\text{sat}}/I_{\text{unsat}}$  uncertainty via error propagation formula.  $R_1$ ,  $R_2$  and  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOEs values have been obtained using the CcpNmr Analysis 2.4 fitting routine<sup>72</sup> and internal scripts. Relaxation experiments have been repeated twice on  $n = 2$  biologically independent samples.

### Isothermal titration calorimetry (ITC)

Proteins and peptides were dialyzed in a Slide-A-lyser mini-dialysis unit with a 2000 MWCO and Biodialyzer with 500 Da MWCO (Harvard Apparatus, US) against 20 mM TrisHCl at pH 7.5, 50 mM NaCl (or 150 mM NaCl when explicitly stated). ITC data were collected on a MicroCal PEAQ-ITC instrument (Malvern). The cell temperature was set to 37 °C, the syringe stirring speed to 750 rpm, and reference power to 10  $\mu\text{cal}/\text{sec}$ . frHMGB1 (frHMGB1-TL, Ac-pep) and CXCL12 were loaded into the cell and syringe at concentrations of -10 and -600  $\mu\text{M}$ , respectively. The MicroCal PeakITC software (Malvern) was applied for initial data analysis. For global fitting, thermograms were integrated using NITPIC 2.0.0<sup>38</sup> and SEDPHAT 15.2b<sup>77</sup>. Data were fit with the two non-symmetric sites microscopic  $K$  model<sup>39</sup> applying Simplex and Marquardt-Levenberg optimization algorithms, as implemented in SEDPHAT 15.2b. Error estimates were based on the covariance matrix generated by the Marquardt-Levenberg algorithm. Experiments have been repeated on  $n = 2$  independent samples.

### Microscale thermophoresis (MST)

MST experiments were performed at 24 °C on a NanoTemper® Monolith NT.115 instrument. Binding between Ac-pep and CXCL12 was monitored titrating CXCL12 (16-points) into 50 nM N-terminal-5,6-FAM-labelled Ac-pep (CASLO ApS, Denmark), using the blue filter, 20% LED power and medium MST power. Binding between frHMGB1, dsHMGB1 or frHMGB1-TL and CXCL12 was monitored titrating the different HMGB1 proteins (16-points) into 50 nM 6His-tagged CXCL12, non-covalently labelled with the NT-647 conjugated tris-NTA (RED-tris-NTA) fluorescence dye, using the red filter, 40% LED power and medium MST power. Before MST titrations the proteins (the ligand and the fluorescently labelled target) were dialyzed against the same buffer, 20 mM  $\text{Na}_2\text{HPO}_4/\text{Na}_2\text{HPO}_4$  pH 7.3, 0.05% TWEEN and 20 mM or 150 mM NaCl. In the case of 5,6-FAM-labelled Ac-pep, the lyophilized peptide was dissolved directly in the MST buffer and the pH adjusted to pH 7.3.

The 16 titration points of each experiment were made through serial dilution of the ligand stock into MST buffer and then addition of a constant amount of fluorescently labelled target (50 nM). Before mixing, both the ligand and the fluorescently labelled target were centrifuged at 15,000 g, 4 °C, for 10 min. Maximum concentrations of frHMGB1, frHMGB1-TL, dsHMGB1 and CXCL12 ligands in the titrations were 104–343  $\mu\text{M}$ , 315  $\mu\text{M}$ , 300  $\mu\text{M}$  and 287–295  $\mu\text{M}$ , respectively. Complex samples were incubated for 30 min before loading into NanoTemper premium capillaries. Each experiment was repeated on  $n = 3$  independent samples, data points are the average of the triplicates and the error bars correspond to the standard deviation.

Since CXCL12 addition induced  $>10\%$  variation in the fluorescence of 5,6-FAM-Ac-Pep, thermophoresis traces could not be used to measure binding affinity, we therefore used the quenching of the 5,6-FAM-Ac-Pep fluorescence upon binding to estimate the  $K_D$ . Data analyses were carried out using NanoTemper Analysis 2.3 software and the  $K_D$  model fitting (one binding site).

### Analytical ultracentrifugation (AUC)

Sedimentation velocity experiments were performed on an Optima XLI (Beckman Coulter) using an A50 Ti eight-hole rotor and with seven 400  $\mu\text{l}$  samples in standard dual-sector Epon centerpieces equipped with sapphire windows. Absorbance data were acquired at 250 and 280 nm simultaneously with the absorbance scanner in the continuous mode with radial increments of 0.003 cm. Three assembled centrifugation cells containing, respectively, free CXCL12 (or CXCL12-LM) (38.2  $\mu\text{M}$  or 37.6  $\mu\text{M}$ ), frHMGB1 (15.6  $\mu\text{M}$ ), frHMGB1-TL (15.6  $\mu\text{M}$ ) and the frHMGB1: CXCL12 (or frHMGB1: CXCL12-LM) mixture at the loading ratios of 1:2.5, 1:5, 1:7.5 (with frHMGB1 or frHMGB1-TL at 7.8  $\mu\text{M}$  concentration), pH 7.5, 20 mM TrisHCl, 50 mM NaCl (or 150 mM NaCl when explicitly stated), were equilibrated at 20 °C under vacuum for approximately 1.5 h prior starting the experiment. Subsequently, centrifugation was performed at 163,000.44 g with 90 scans. The highest protein concentration was determined by the absorbance for which the linear relationship according to Lambert-Beer Law was still guaranteed (O.D. max = 1.0). The buffer density and viscosity were estimated using SEDNTERP 1.0<sup>78</sup>.

The sedimentation coefficient distributions ( $c(s)$ ) at 280 nm of the single proteins were obtained by applying the diffusion-deconvoluted  $c(s)$  model, implemented in SEDFIT 16.1c<sup>41</sup>. Concentration profiles in terms of absorbance ( $a(r,t)$ ) were modelled as the sum of Lamm Equation solutions scaled by a continuous distribution  $c(s)$  as follows:

$$a(r,t) \cong \int_{s_{\text{min}}}^{s_{\text{max}}} c_k(s) \chi(s, D_k(s), r, t) ds \quad (2)$$

where  $s$  is the sedimentation coefficient,  $\chi(s, D(s), r, t)$  is the Lamm Equation solution that is dependent on  $D(s)$ , the corresponding diffusion coefficient,  $r$ , radius from the center of rotation, and  $t$ , the time from the beginning of the experiment.

Multisignal sedimentation velocity (MSSV) analysis<sup>44</sup> was performed to determine the stoichiometry of the complex formed by frHMGB1 and CXCL12 (or CXCL12-LM). It is important to note that in MSSV, like in  $c(s)$  model, the data are not fitted a priori to a binding stoichiometry model (e.g. 1:1 or 1:2) and no binding model assumption is made. The experimental sedimentation data, whose physical observable is absorbance, are fitted with mathematical equations (the Lamm equations) describing the sedimentation process without assuming any binding stoichiometry. The sedimentation process is described as a superposition of normalized Lamm equation solutions of ideally sedimenting species at a range of sedimentation coefficients  $s(1...N)$ , with  $N$  indicating the range of molecules size expressed as sedimentation coefficient in Svedberg and using the hydrodynamic scaling law to estimate the corresponding diffusion coefficients. The Lamm equation solutions once superimposed give the number of species, molecular weight and weight averaged shape. More in detail, in MSSV the standard  $c(s)$  approach is modified to deconvolute the contributions of individual species in a component distribution  $c_k(s)$  where  $k$  represents the individual components of a mixture. The absorbance at wavelength  $\lambda$  ( $a_\lambda(r,t)$ ) is modelled as:

$$a(r,t) \cong \sum_{k=1}^K \epsilon_\lambda^k \int_{s_{\min}}^{s_{\max}} c(s) \chi(s, D(s), r, t) ds \quad (3)$$

where  $l$  is the path length,  $k$  is the number of solutes present, and  $c_k(s)$  is a continuous distribution for component  $k$ .

MSSV deconvolution is possible when the complex components have sufficiently different spectral properties<sup>44</sup>, i.e.

$$D_{\text{norm}} = \frac{\|\det \epsilon_k^\lambda\|}{\prod_k \|\epsilon_k\|} > 0.065 \quad (4)$$

As the molar extinction coefficients of frHMGB1 ( $\epsilon_{280} = 20872.8 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\epsilon_{250} = 7987.7 \text{ M}^{-1} \text{ cm}^{-1}$ ) and CXCL12 ( $\epsilon_{280} = 9907.2 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\epsilon_{250} = 4704.2 \text{ M}^{-1} \text{ cm}^{-1}$ ) or CXCL12-LM ( $\epsilon_{280} = 8413.4 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\epsilon_{250} = 4840.5 \text{ M}^{-1} \text{ cm}^{-1}$ ) delivered sufficient spectral discrimination with  $D_{\text{norm}} > 0.08$  (for frHMGB1:CXCL12) and  $D_{\text{norm}} > 0.16$  (for frHMGB1:CXCL12-LM), it was possible to use SEDPHAT 15.2b to perform global multi-signal analysis of the sedimentation boundary associated to the co-sedimenting complex<sup>44</sup>. MSSV of the frHMGB1:CXCL12 (7.7  $\mu\text{M}$ :43  $\mu\text{M}$ ) and frHMGB1:CXCL12-LM (6.9  $\mu\text{M}$ :51  $\mu\text{M}$ ) heterocomplexes were collected at 250 nm and 280 nm and globally fitted using the multi-wavelength discrete/continuous distribution analysis with mass constraints in SEDPHAT 15.2b<sup>44</sup>. Integration of the resulting  $c_k(s)$  distributions revealed the content of each protein component under the peak at a given sedimentation value. Plots of the signal profiles, fits, residuals and the MSSV results were generated using GUSI 1.4.2<sup>79</sup>.

### Small angle X-ray scattering (SAXS)

The experiments were performed at the ESRF bioSAXS beamline BM29, Grenoble, France at a detector distance of 2.869 m. CXCL12 and frHMGB1 were measured in batch mode at 20 °C using the sample changer immediately after protein thawing and centrifugation (30 min at 16,000 g); 45  $\mu\text{L}$  of sample solution at three different concentrations (1.95, 3.0 and 3.9 mg/mL per each protein, 20 mM Tris pH 7.5, 50 mM NaCl) were used. To characterize the frHMGB1:CXCL12 complex we tested the following protein molar ratio conditions: 1:1, 1:2, 1:4 and 1:6 (frHMGB1:CXCL12, with frHMGB1 1.95 mg/mL). After incubation and centrifugation, the supernatants were immediately measured in the SAXS beamline.  $n=10$  frames of 0.5 s each were collected for each sample (frHMGB1:CXCL12, free components at different protein concentration). Data from the frHMGB1 and CXCL12 dilutions were merged following standard procedures to create an idealized scattering curve, using PRIMUS within ATSAS 3.2.1<sup>80</sup>. The pair distribution function  $P(r)$  was calculated using GNOM version 5.0 (r14886)<sup>81</sup>.

Protein molecular masses were estimated using both Porod volume and scattering mass contrast methods. All the plots were generated using OriginPro (Version 2022, OriginLab Corporation, Northampton, MA, USA).

To model the frHMGB1:CXCL12 heterocomplex, we employed a multistep approach relying on complementary experimental knowledge (NMR and AUC). The SAXS curve of frHMGB1 in the presence of two equivalents of CXCL12 was used, as in this condition frHMGB1 is almost fully saturated and the contribution of free CXCL12 to the scattering signal is negligible, as also confirmed by OLIGOMER analysis (within ATSAS 3.2.1<sup>80</sup>) (Supplementary Methods and Supplementary Fig. 15). This SAXS curve was used to generate with EOM a representative frHMGB1 starting structure for subsequent rigid docking on CXCL12 (pdb: 2KEE). In particular, two initial rigid docking models were generated. The first one was obtained with SASREF<sup>47</sup> (within ATSAS 3.2.1<sup>80</sup>) combining the solution scattering data and guiding the docking of CXCL12 (residues 23–28 and 66–67) on frHMGB1 (residues 15–45), as suggested by NMR experiments. The second model was obtained docking CXCL12 (residues 23–28 and 66–67) onto HMGB1 acidic IDR performing a global search with FoXSDock version main.ec6dbc2<sup>48</sup> (Supplementary Fig. 13). Next, to describe the dynamic and fuzzy nature of the heterocomplex, we fixed the protein-protein interaction surfaces obtained with SASREF<sup>47</sup> and FoXSDock version main.ec6dbc2<sup>48</sup>, and allowed the rest of frHMGB1 to explore a wide range of conformations using the EOM algorithm within ATSAS 3.2.1<sup>80</sup>. Similarly, frHMGB1 alone was modelled using the EOM approach (parameters for the analysis are in Supplementary Table 4). The missing residues connecting the rigid bodies and the modelled segments were added with MODELLER 10.1<sup>82</sup>. The  $\chi^2$  values of EOM fits over the 0.1–3 nm<sup>-1</sup>  $q$  range of experimental SAXS curves were determined with CORMAP within ATSAS 3.2.1<sup>80</sup>.

All SAXS data were deposited into SASBDB data bank (CXCL12, SASDB ID: SASDRG9); frHMGB1, SASDB ID: SASDRH9; frHMGB1:CXCL12, SASDB ID: SASDRJ9.

### Molecular images

Molecular images were generated by PyMOL Molecular Graphics System, open source version, Schrödinger, LLC and UCSF Chimera 1.16<sup>83</sup> and assembled with Inkscape 0.92.

### Cell line and treatments

AB1 mouse malignant mesothelioma cells (MM; Cell Bank, Australia) were cultured in RPMI 1640 (Life Technologies, UK), supplemented with 5% v/v fetal bovine serum (FBS; Life Technologies, UK), 2 mM L-glutamine, and 100 U/ml penicillin/streptomycin at 37 °C; 5% CO<sub>2</sub>. Cells were not cultured past passage 10 after cell thawing. To show binding of the HMGB1:CXCL12 heterocomplex to CXCR4, cells were incubated or not at 37 °C with increasing concentrations of AMD3100 and then fixed and quantified for the PLA signal. To validate these observations, we generated a *Cxcr4* knockout AB1 cell line (see further) to determine whether CXCR4 is indeed required for frHMGB1:CXCL12 complex formation. Untreated *Cxcr4*<sup>-/-</sup> and wild-type AB1 cells were fixed and quantified for the PLA signal. As a further verification, wild-type and *Cxcr4*<sup>-/-</sup> cells were co-cultured at a 3:1 and a 1:3 ratio, fixed, and quantified for the PLA signal to demonstrate that the ratio of cells with high and low PLA signal would be 3:1 (or 1:3), confirming the previous experiment (Supplementary Fig. 16).

To show the effect of Ac-pep on the binding of the frHMGB1:CXCL12 heterocomplex to the receptor, cells were incubated or not at 37 °C with increasing concentrations of Ac-pep and then fixed and quantified for the PLA signal.

To measure the binding of the frHMGB1:CXCL12 heterocomplex to its receptor, cells were washed with 10 mM Tris-HCl, pH 5.3, washed with RPMI and then incubated for 20 min at 4 °C with RPMI containing the indicated concentrations of preformed full-length or tailless

frHMGB1•CXCL12 heterocomplexes. Cells were then fixed and quantified for the PLA signal.

### Proximity ligation assay (PLA)

AB1 cells ( $2 \times 10^4$ ) were seeded onto 15 mm glass coverslips and incubated overnight at 37 °C/5% CO<sub>2</sub>, and treated the following day as described above. To perform PLA, cells were fixed for 10 min in 4% paraformaldehyde in PHEM buffer (1:1) at room temperature (RT), washed in 0.2% bovine serum albumin (BSA)/PBS and blocked with 10% goat serum in 4% BSA/PBS for 1 h at RT. Cells were then incubated overnight at 4 °C with two primary antibodies: mouse monoclonal anti-HMGB1 (1:1000; HMGBiotech, HM-901) and goat polyclonal anti-CXCL12 (1:50; R&D Systems, #AF-310-NA). Following three washes with 0.2% BSA/PBS, cells were incubated with secondary oligonucleotide-linked antibodies for 1 h at 37 °C (anti-mouse PLUS [#DUO92001] and anti-goat MINUS [#DUO92006], Duolink, Sigma-Aldrich; 1:5 dilution in Duolink antibody diluent [DUO82008]) and processed according to the Duolink PLA fluorescence protocol (Sigma). Finally, cells were stained with Phalloidin-FITC (1:500; Sigma-Aldrich, #P5282) and Hoechst (1 µg/mL; Sigma-Aldrich, #33342), and mounted on microscope slides with FluorSave reagent (Merck, #345789).

All fields of view (FOV) for PLA were acquired using a Leica TCS SP5 X confocal microscope (Leica Biosystems) with a 63x objective, using channels for Hoechst (405 nm), FITC (488 nm), and Texas Red (561 nm). Z-step size was set at 0.69 µm and the top and bottom of the cells were ascertained manually prior to acquiring each image. Each image was stacked to max intensity using ImageJ/Fiji software and saved as tiff files. These images were then used to create cytoplasmic masks using Cellpose<sup>84</sup>. Python 3.10.0 was used to run Cellpose. Cytoplasmic masks were used to ascertain cytoplasmic specific PLA signal in ImageJ/Fiji. Data are expressed as either as the average number of PLA dots per cell (Fig. 8b) or as Raw Integrated Density (sum of all pixel intensities in region of interest, in arbitrary units) normalized to the area of the respective cytosolic regions (Fig. 8a, c where the number of dots was difficult to count due to overlaps). All PLA experiments have been repeated twice. GraphPad Prism, version 8.4.0 was used for the generation of original graphs and statistical tests.

### Generation of *Cxcr4* knockout AB1 cells

*Cxcr4* knockout AB1 cells were generated using the LentiCRISPRv2 *BsmBI* CRISPR/Cas9 system to obtain two cuts flanking the last exon of mouse *Cxcr4* gene. Guide RNAs were designed using the UCSC Genome Browser and Primer3 program (Sg1FWD: TGTGGTTATGCTGTGTG; Sg1REV: ACAAAACCAATACGACACACT; Sg2FWD: TGAATGACGCTTTTCATCC; Sg2REV: ACTTTACCTGCAAAAGTAGG). The LentiCRISPRv2 vector was digested using *BsmBI* (#R0739; New England Biolabs), ligated with the oligonucleotide guides and transformed into Stbl3 bacterial cells, which were plated on ampicillin plates and screened for the correct CXCR4-sgRNA construct. Vector lentiviruses were then produced in HEK293 cells after transfection with lipofectamine 3000 of VSV-G envelope (#14888; Addgene), ps-PAX2 packaging (#12260; Addgene) and CXCR4-sg-RNA plasmids. AB1-B/c-LUC cells were then infected with this lentivirus and selected using Neomycin (10 µg/mL); single cells were plated in a 96-well plate and allowed to grow to confluence before DNA extraction and PCR with primers flanking the deletion. As a further verification, FACS analysis was carried out to determine CXCR4 expression (Supplementary Fig. S17).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All SAXS data have been deposited to the SASBDB with the following codes: SASDRH9; SASDRG9; SASDRJ9. NMR backbone assignments for

CXCL12-LM used in this study are available under BMRB entry ID 52209. The experimental data that support the findings of this study are shown in the article and its supplementary materials. The Source data underlying all Figures and Supplementary are provided as Source Data file with this paper and are available at [<https://doi.org/10.6084/m9.figshare.24574522>]. All other data are available from the authors upon request. NMR backbone assignments of HMGB1, HMGB1-TL, CXCL12 are available under BMRB entry: ID 15149; ID 15148 and ID 16519, respectively. The NMR Coordinates of CXCL12, CXCL12-LM, and HMGB1-TL are available in the PDB with the following accession codes 2KEE; 2N55; 2YRQ, respectively. The crystallographic structure of CXCL12 with two monomers in the asymmetric unit is available in the PDB with the code 1QG7. The AlphaFold model of HMGB1 is available in the AlphaFold Protein Structure Database with the following entry AF-P63159. Source data are provided with this paper.

### References

- Zlotnik, A. & Yoshie, O. The chemokine superfamily revisited. *Immunity* **36**, 705–716 (2012).
- Hughes, C. E. & Nibbs, R. J. B. A guide to chemokines and their receptors. *FEBS J.* **285**, 2944–2971 (2018).
- Drurya, L. J. et al. Monomeric and dimeric CXCL12 inhibit metastasis through distinct CXCR4 interactions and signaling pathways. *Proc. Natl Acad. Sci. USA* **108**, 17655–17660 (2011).
- Miller, M. C. & Mayo, K. H. Chemokines from a structural perspective. *Int. J. Mol. Sci.* **18**, 2088 (2017).
- Veldkamp, C. T., Peterson, F. C., Pelzek, A. J. & Volkman, B. F. The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci.* **14**, 1071–1081 (2005).
- Proudfoot, A. E. I. & Uguccioni, M. Modulation of chemokine responses: synergy and cooperativity. *Front. Immunol.* **7**, 183 (2016).
- Cecchinato, V., D'Agostino, G., Raeli, L. & Uguccioni, M. Chemokine interaction with synergy-inducing molecules: fine tuning modulation of cell trafficking. *J. Leukoc. Biol.* **99**, 851–855 (2016).
- Gouwy, M., Schiraldi, M., Struyf, S., Van Damme, J. & Uguccioni, M. Possible mechanisms involved in chemokine synergy fine tuning the inflammatory response. *Immunol. Lett.* **145**, 10–14 (2012).
- Weber, C. & Koenen, R. R. Fine-tuning leukocyte responses: towards a chemokine 'interactome'. *Trends Immunol.* **27**, 268–273 (2006).
- Paoletti, S. et al. A rich chemokine environment strongly enhances leukocyte migration and activities. *Blood* **105**, 3405–3412 (2005).
- Koenen, R. R. et al. Disrupting functional interactions between platelet chemokines inhibits atherosclerosis in hyperlipidemic mice. *Nat. Med.* **15**, 97–103 (2009).
- Venez, D. et al. Perivascular expression of CXCL9 and CXCL12 in primary central nervous system lymphoma: T-cell infiltration and positioning of malignant B cells. *Int. J. Cancer* **127**, 2300–2312 (2010).
- von Hundelshausen, P. et al. Chemokine interactome mapping enables tailored intervention in acute and chronic inflammation. *Sci. Transl. Med.* **9**, eaah6650 (2017).
- Eckardt, V. et al. Chemokines and galectins form heterodimers to modulate inflammation. *EMBO Rep.* **21**, e47852 (2020).
- Schiraldi, M. et al. HMGB1 promotes recruitment of inflammatory cells to damaged tissues by forming a complex with CXCL12 and signaling via CXCR4. *J. Exp. Med.* **209**, 551–563 (2012).
- Bianchi, M. E. & Mezzapelle, R. The chemokine receptor CXCR4 in cell proliferation and tissue regeneration. *Front. Immunol.* **11**, 2109 (2020).
- Scaffidi, P., Misteli, T. & Bianchi, M. E. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature* **418**, 191–195 (2002).
- Venereau, E., Ceriotti, C. & Bianchi, M. E. DAMPs from cell death to new life. *Front. Immunol.* **6**, 422 (2015).

19. Kang, R. et al. HMGB1 in health and disease. *Mol. Asp. Med.* **40**, 1–116 (1998).
20. Venereau, E. et al. Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release. *J. Exp. Med.* **209**, 1519–1528 (2012).
21. Bianchi, M. E. et al. High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair. *Immunol. Rev.* **280**, 74–82 (2017).
22. D’Agostino, G. et al.  $\beta$ -Arrestin1 and  $\beta$ -Arrestin2 are required to support the activity of the CXCL12/HMGB1 heterocomplex on CXCR4. *Front. Immunol.* **11**, 550824 (2020).
23. Mezzapelle, R. et al. CXCR4 engagement triggers CD47 internalization and antitumor immunization in a mouse model of mesothelioma. *EMBO Mol. Med.* **13**, e12344 (2021).
24. De Leo, F. et al. Diflunisal targets the HMGB1/CXCL12 heterocomplex and blocks immune cell recruitment. *EMBO Rep.* **20**, e47788 (2019).
25. Veldkamp, C. T., Peterson, F. C., Pelzek, A. J. & Volkman, B. F. The monomer – dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci.* **4**, 1071–1081 (2005).
26. Wang, X. et al. Dynamic autoinhibition of the HMGB1 protein via electrostatic fuzzy interactions of intrinsically disordered. *Reg. J. Mol. Biol.* **433**, 167122 (2021).
27. Stott, K., Watson, M., Howe, F. S., Grossmann, J. G. & Thomas, J. O. Tail-mediated collapse of HMGB1 is dynamic and occurs via differential binding of the acidic tail to the A and B domains. *J. Mol. Biol.* **403**, 706–722 (2010).
28. Knapp, S. et al. The long acidic tail of high mobility group Box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes. *Biochemistry* **43**, 11992–11997 (2004).
29. Watson, M., Stott, K. & Thomas, J. O. Mapping intramolecular interactions between domains in HMGB1 using a Tail-truncation. *Approach J. Mol. Biol.* **374**, 1286–1297 (2007).
30. Fassi, E. M. A. et al. Oxidation state dependent conformational changes of HMGB1 regulate the formation of the CXCL12/HMGB1 heterocomplex. *Comput. Struct. Biotechnol. J.* **17**, 886–894 (2019).
31. Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).
32. Ahmed, R. & Forman-Kay, J. D. NMR insights into dynamic, multi-valent interactions of intrinsically disordered regions: from discrete complexes to condensates. *Essays Biochem.* **66**, 863–873 (2022).
33. Waudby, C. A., Ramos, A., Cabrita, L. D., Christodoulou, J. & Two-Dimensional, N. M. R. Lineshape analysis. *Sci. Rep.* **6**, 24826 (2016).
34. Ziarek, J. J. et al. Sulfopeptide probes of the CXCR4/CXCL12 interface reveal oligomer-specific contacts and chemokine allostery. *ACS Chem. Biol.* **8**, 1955–1963 (2013).
35. Veldkamp, C. T. et al. Monomeric structure of the cardioprotective chemokine SDF-1/CXCL12. *Protein Sci.* **18**, 1359–1369 (2009).
36. Mittag, T. et al. Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc. Natl Acad. Sci. USA* **105**, 17772–17777 (2008).
37. Borgia, A. et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 (2018).
38. Keller, S. et al. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal. Chem.* **84**, 5066–5073 (2012).
39. Brautigam, C. A. Fitting two- and three-site binding models to isothermal titration calorimetric data. *Methods* **76**, 124–136 (2015).
40. Mensah, M. A. et al. Aberrant phase separation and nucleolar dysfunction in rare genetic diseases. *Nature* **614**, 564–571 (2023).
41. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
42. Balbo, A. & Schuck, P. Analytical ultracentrifugation in the study of protein self-association and heterogeneous protein-protein interactions. *Golemis, E. Adams, P. D. Protein-Protein Interact. Cold Spring Harb. Lab. Press. Cold Spring Harb. New York* 253–277 (2005).
43. Brautigam, C. A., Padrick, S. B. & Schuck, P. Multi-signal sedimentation velocity analysis with mass conservation for determining the stoichiometry of protein complexes. *PLoS One* **8**, e62694 (2013).
44. Padrick, S. B. & Brautigam, C. A. Evaluating the stoichiometry of macromolecular complexes using multisignal sedimentation velocity. *Methods* **54**, 39–55 (2011).
45. Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).
46. Tuukkanen, A. T. & Svergun, D. I. Weak protein-ligand interactions studied by small-angle X-ray scattering. *FEBS J.* **281**, 1974–1987 (2014).
47. Petoukhov, M. V. & Svergun, D. I. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**, 1237–1250 (2005).
48. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.* **44**, W424–W429 (2016).
49. Mezzapelle, R. et al. Human malignant mesothelioma is recapitulated in immunocompetent BALB/c mice injected with murine AB cells. *Sci. Rep.* **6**, 22850 (2016).
50. Bugge, K. et al. Interactions by disorder – A matter of context. *Front. Mol. Biosci.* **7**, 110 (2020).
51. Uversky, V. N. Intrinsically disordered proteins and their “mysterious” (Meta) physics. *Front. Phys.* **7**, 10 (2019).
52. Arbesú, M., Iruela, G., Fuentes, H., Teixeira, J. M. C. & Pons, M. Intramolecular fuzzy interactions involving intrinsically disordered domains. *Front. Mol. Biosci.* **5**, 39 (2018).
53. Fuxreiter, M. Electrostatics tunes protein interactions to context. *Proc. Natl Acad. Sci. USA* **119**, e2209201119 (2022).
54. Holehouse, A. S. & Kragelund, B. B. The molecular basis for cellular function of intrinsically disordered protein regions. *Nat. Rev. Mol. Cell Biol.* <https://doi.org/10.1038/s41580-023-00673-0> (2023). Online ahead of print
55. Miskei, M. et al. Fuzziness enables context dependence of protein interactions. *FEBS Lett.* **591**, 2682–2695 (2017).
56. Wang, Q., Zeng, M., Wang, W. & Tang, J. The HMGB1 acidic tail regulates HMGB1 DNA binding specificity by a unique mechanism. *Biochem. Biophys. Res. Commun.* **360**, 14–19 (2007).
57. Cato, L., Stott, K., Watson, M. & Thomas, J. O. The interaction of HMGB1 and linker histones occurs through their acidic and basic tails. *J. Mol. Biol.* **384**, 1262–1272 (2008).
58. Watson, M., Stott, K., Fischl, H., Cato, L. & Thomas, J. O. Characterization of the interaction between HMGB1 and H3-a possible means of positioning HMGB1 in chromatin. *Nucleic Acids Res.* **42**, 848–859 (2014).
59. Rowell, J. P., Simpson, K. L., Stott, K., Watson, M. & Thomas, J. O. HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure* **20**, 2014–2024 (2012).
60. Ziarek, J. J. et al. Heparin oligosaccharides inhibit chemokine (CXC Motif) Ligand 12 (CXCL12) cardioprotection by binding orthogonal to the dimerization interface, promoting oligomerization, and competing with the chemokine (CXC Motif) Receptor 4 (CXCR4) N Terminus. *J. Biol. Chem.* **288**, 737–746 (2013).
61. Sharma, R., Raduly, Z., Miskei, M. & Fuxreiter, M. Fuzzy complexes: Specific binding without complete folding. *FEBS Lett.* **589**, 2533–2542 (2015).
62. Uversky, V. N. A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Sci.* **22**, 693–724 (2013).



63. Borg, M. et al. Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc. Natl Acad. Sci. USA* **104**, 9650–9655 (2007).
64. Bianchi, M. E. HMGB1 loves company. *J. Leukoc. Biol.* **86**, 573–576 (2009).
65. Agresti, A., Scaffidi, P., Riva, A., Caiola, V. R. & Bianchi, M. E. GR and HMGB1 interact only within chromatin and influence each other's residence time. *Mol. Cell* **18**, 109–121 (2005).
66. Celona, B. et al. Substantial histone reduction modulates genome-wide nucleosomal occupancy and global transcriptional output. *PLoS Biol.* **9**, e1001086 (2011).
67. Agresti, A. & Bianchi, M. E. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* **13**, 170–178 (2003).
68. Clouser, A. F. et al. Interplay of disordered and ordered regions of a human small heat shock protein yields an ensemble of 'quasi-ordered' states. *Elife* **8**, e50259 (2019).
69. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).
70. Tsafo, K., Tiwari, P. B., Forman-Kay, J. D., Metallo, S. J. & Toretzky, J. A. Targeting intrinsically disordered transcription factors: changing the paradigm. *J. Mol. Biol.* **430**, 2321–2341 (2018).
71. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
72. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins Struct. Funct. Genet.* **59**, 687–696 (2005).
73. Grzesiek, S., Stahl, S. J., Wingfield, P. T. & Bax, A. The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry* **35**, 10256–10261 (1996).
74. Schumann, F. H. et al. Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions. *J. Biomol. NMR* **39**, 275–289 (2007).
75. Farrow, N. A. et al. Backbone dynamics of a free and phosphopeptide-complexed Src- homology 2 domain studied by <sup>15</sup>N relaxation. *Biochemistry* **33**, 5984–6003 (1994).
76. Yip, G. N. B. & Zuiderweg, E. R. P. Improvement of duty-cycle heating compensation in NMR spin relaxation experiments. *J. Magn. Reson.* **176**, 171–178 (2005).
77. Houtman, J. C. D. et al. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling. *Protein Sci.* **16**, 30–42 (2007).
78. Philo, J. S. SEDNTERP: a calculation and database utility to aid interpretation of analytical ultracentrifugation and light scattering data. *Eur. Biophys. J.* **52**, 233–266 (2023).
79. Brautigam, C. A. Calculations and publication-quality illustrations for analytical ultracentrifugation data. *Methods Enzymol.* **562**, 109–133 (2015).
80. Manalastas-Cantos, K. et al. ATSAS 3.0: Expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
81. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
82. Eswar, N. et al. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **Chapter 2**, Unit 2.9 (2007).
83. Pettersen, E. F. et al. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
84. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
85. Ngo, T. et al. Crosslinking-guided geometry of a complete CXC receptor-chemokine complex and the basis of chemokine subfamily selectivity. *PLoS Biol.* **18**, e3000656 (2020).

## Acknowledgements

The research reported here has received funding from AIRC (Associazione Italiana per la Ricerca sul Cancro) under IG 2018 - ID. 21440 project - P.I. GM, and IG 2020 - ID. 24702 P.I. MEB, and MFAG27415 P.I. RM; by Fondazione Buzzi Unicem - P.I. MEB. MVM was supported by a FIRC-AIRC fellowship 24118 for Italy. LSC was supported by the RENOIR European Training Network. CZ was supported by FSHD Global Research Foundation FSHD-Winter2021-5008608933, GG has received funding from the University of Padova, Starting Grant STARS@UNIPD. The authors acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities under proposal number MX2386 and thank Mark Tully and Petra Pernot for assistance and support in using beamline BM29. We wish to thank Dr Joy Zhao, (Laboratory of Dynamics of Macromolecular Assembly National Institute of Biomedical Imaging and Bioengineering) for useful discussions on AUC data. Part of the present work was performed by MVM in fulfillment of the requirements for obtaining a Ph.D. degree at Vita-Salute San Raffaele University, Milan, Italy.

## Author contributions

M.V.M., C.C., C.P. produced recombinant proteins. M.V.M., F.D.L., M.G. and G.Q. analyzed NMR experiments. M.V.M., T.S. and S.R. performed and analyzed ITC experiments. M.V.M. performed and analyzed AUC experiments and prepared a draft of the manuscript. G.Q. and M.V.M. performed the NMR experiments. F.D.L. and C.Z. performed and analyzed MST experiments. L.S.C., F.D.M. and M.C. performed and analyzed PLA experiments and generated Cxcr4 KO AB1 cells. F.C. performed flow cytometry. R.M. participated to the design of cell-based experiments. M.E.B. coordinated all experiments with cells and contributed to the writing of the manuscript. G.G. supervised the SAXS experiments and performed all the SAXS data analysis with the support of M.V.M., F.D.L. and M.G. M.G. generated the model of the interaction, participated to the interpretation of all the data and in manuscript preparation. G.M. supervised the study and was involved in all aspects of the experimental design, data analysis, and manuscript preparation. All authors critically reviewed the text and figures.

## Competing interests

The authors declare the following competing interests: L.S.C. is an employee and M.E.B. is founder and part-owner of HMGBiotech, a company that provides goods and services related to HMGB proteins. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains Supplementary Material available at <https://doi.org/10.1038/s41467-024-45505-7>.

**Correspondence** and requests for materials should be addressed to Michela Ghitti or Giovanna Musco.

**Peer review information** *Nature Communications* thanks Junji Iwahara, Erik Martin, Theodora Myrto Perdikari and the other, anonymous, reviewer for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024